

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 13:53:21 ; Search time 167 Seconds
(without alignments)
1199.652 Million cell updates/sec

Title: US-09-815-379-4

Perfect score: 2789
Sequence: 1 MGSPAAPGALGYVRETRH.....RQKGAATKVKHYHLGGP 518

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2789	100.0	518	4	AAE11887
2	2556	91.6	480	6	AAO16438
3	2541.5	91.1	479	6	ABU03467
4	2541.5	91.1	479	7	ABU04323
5	2541.5	91.1	479	7	ADN38698
6	2464	88.3	480	7	ABU64322
7	2414	86.6	458	8	ABM84710
8	1138	40.8	707	5	ABBS7289
9	1128.5	40.5	706	8	ADRI4065
10	1126.5	40.4	706	7	ADJ70182
11	1122.5	40.2	706	3	AAJ78793
12	1122.5	40.2	706	4	AAJ78793
13	1122.5	40.2	706	8	ADL82847
14	1122.5	40.2	706	8	ADRI4017
15	1095.5	39.3	706	2	AAJ68743
16	1095.5	39.3	706	3	AAJ78792
17	582	20.9	810	3	AAJ73351
18	580	20.8	667	8	ADP22538
19	576.5	20.7	803	2	AAW81756
20	575.5	20.6	803	7	ADD45510
21	575.5	20.6	803	8	ADG30638
22	574.5	20.6	804	4	AAJ39272
23	574.5	20.6	804	4	AAJ39272
24	568.5	20.4	756	7	ADFI7459
25	543.5	19.5	765	6	ABP96229

ALIGNMENTS

RESULT 1

AAE11887

ID AAE11887 standard; protein; 518 AA.

XX AC AAE11887;

XX DT 18-DEC-2001 (first entry)

XX DE Angiogenesis associated human BAZF (hBAZF) protein.

XX KW Angiogenesis associated protein; AAP; cytostatic; cardiant; gene therapy; ophthalmological; vulnary; myocardial infarction; macular degeneration; diabetic retinopathy; angiogenesis; wound healing; prophylactic; vaccine; rheumatoid arthritis; psoriasis; drug screening; tumour; transplantation; cancer; therapeutic; diagnostic; human; BAZF protein.

XX OS Homo sapiens.

XX PN WO200170808-A2.

XX PD 27-SEP-2001.

XX PF 22-MAR-2001; 2001WO-US009609.

XX PA (CURA-) CURAGEN CORP.

XX PA (GETH) GENENTECH INC.

XX PI Rastelli LK, Gerritsen M;

XX DR WPI: 2001-602775/68.

XX N-PSDB; AAD19116.

XX Novel angiogenesis associated polypeptides and polynucleotides encoding the polypeptides, useful for modulating angiogenesis and for treating tumors and cancers.

XX PS Claim 1; Page 12-13; 159pp; English.

XX CC The invention relates to angiogenesis associated proteins (AAP) and their corresponding cDNA molecules, which are useful for modulating angiogenesis. AAP proteins and nucleic acids are useful for promoting wound healing, for example after organ transplantation, and in the treatment of tumours, myocardial infarction, cancers, diabetic retinopathy, macular degeneration, psoriasis and rheumatoid arthritis. AAP proteins and DNA's are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders including

26 539.5 19.3 765 8 ADR09346 ADR09346 Human pro
27 509.5 18.3 539 7 ADG14300 ADG14300 Human HCK
28 485.5 17.4 545 8 ADR08909 ADR08909 Human pro
29 483.5 17.3 697 7 ADJ70450 ADJ70450 Human hea
30 483 17.3 610 5 AAM47790 AAM47790 Murine my
31 482.5 17.3 611 4 AAM40035 AAM40035 Human pol
32 482.5 17.3 697 8 ADQ21588 ADQ21588 Human sof
33 482 17.3 610 4 AAB93761 AAB93761 Human tra
34 482 17.3 610 4 AAB50159 AAB50159 Human tra
35 482 17.3 610 5 AAM47779 AAM47779 Human myo
36 480 17.2 603 4 AAM41821 AAM41821 Human pol
37 479 17.2 775 7 ADC31774 ADC31774 Human nov
38 476 17.1 688 5 ABB82513 ABB82513 Human rep
39 469.5 16.8 682 8 ADM90925 ADM90925 Lung-spec
40 469.5 16.8 711 5 AAO21779 AAO21779 Lung-spec
41 457.5 16.4 641 4 AAB84369 AAB84369 Amino aci
42 457.5 16.4 641 5 ABG61930 ABG61930 Prostata
43 447 16.0 687 7 ADJ70099 ADJ70099 Human hea
44 438.5 15.7 620 4 AAU30585 AAU30585 Novel hum
45 437 645 4 ABG28285 ABG28285 Novel hum

CC those related to angiogenesis, and also in diagnostic applications. AAP
 CC cDNA is also useful in gene therapy. The invention also relates to a
 CC method for screening a tissue sample for tumorigenic potential. AAP
 CC proteins are used to screen drugs or compounds that modulate AAP activity
 CC or expression as well as treating disorders characterised by insufficient
 CC or excessive production of AAP or production of AAP forms that have
 CC decreased or aberrant activity compared to the wild type protein, or
 CC modulate biological function that involve AAP. The present sequence is
 CC human BAZP (hBAZP) protein which is an angiogenesis associated protein
 CC (AAP) of the invention. BAZP is a Bcl-6 (LAZ3) homolog, a transcription
 CC repressor that controls germinal center formation and the T-cell
 CC dependent immune response. Human BAZP plays a role in cell proliferation
 XX
 XX Sequence 518 AA;

Query Match 100.0%; Score 2789; DB 4; Length 518;
 Best Local Similarity 100.0%; Pred. No. 4.7e-191; Indels 0; Gaps 0;
 Matches 518; Conservative 0; Mismatches 0;

QY 1 MGSPAPEGALGYVRETRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACS 60
 DB 1 MGSPAPEGALGYVRETRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACS 60
 QY 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYQLM 120
 DB 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYQLM 120
 QY 121 EHVVOACHRFIOASVEPLGISLRPLEAEPTPTAPPGRSRRSEGHDPPTESRSCSQ 180
 DB 121 EHVVOACHRFIOASVEPLGISLRPLEAEPTPTAPPGRSRRSEGHDPPTESRSCSQ 180
 QY 181 PPSASPDPKACNWKYKIVLNSQAQSLVGRSSGQPCQARLPSCGDEASSSSSS 240
 DB 181 PPSASPDPKACNWKYKIVLNSQAQSLVGRSSGQPCQARLPSCGDEASSSSSS 240
 QY 241 SSSSSEEGPIPGQSRISPTAATVQKCGAPASTPYLLTQAQDTSGSPSERARPLPGSE 300
 DB 241 SSSSSEEGPIPGQSRISPTAATVQKCGAPASTPYLLTQAQDTSGSPSERARPLPGSE 300
 QY 301 PFCQNCVAVAGSSGLDLSVPGDEDPKCYCOLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 DB 301 PFCQNCVAVAGSSGLDLSVPGDEDPKCYCOLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 QY 361 ICGARFNRPANLKTSHRIHSGEKPKYKCTCGSRFQVQVRSPPSGFQKQARGVGQKGF 420
 DB 361 ICGARFNRPANLKTSHRIHSGEKPKYKCTCGSRFQVQVRSPPSGFQKQARGVGQKGF 420
 QY 421 CSSQQRDLKSPPSQVAHLRAHVLHTGKPYKPCPTCTGTRFRHLQTLKSHVRIHTGKPYH 480
 DB 421 CSSQQRDLKSPPSQVAHLRAHVLHTGKPYKPCPTCTGTRFRHLQTLKSHVRIHTGKPYH 480
 QY 481 CDPGLHFRHKSQRLRLHROKHGAATNTKVHYHLGGP 518
 DB 481 CDPGLHFRHKSQRLRLHROKHGAATNTKVHYHLGGP 518

RESULT 2
 AAO16438
 ID AAO16438 standard; protein; 480 AA.
 XX
 AC AAO16438;
 XX
 DT 10-APR-2003 (first entry)
 XX
 DE Human nucleic acid-associated protein (NAAP) - SEQ ID No 35.
 XX
 KW Human; nucleic acid-associated protein; NAAP; arteriosclerosis;
 KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis; AIDS;
 KW cancer; developmental disorder; renal tubular acidosis; anaemia; asthma;
 KW mental retardation; neurological disorder; Alzheimer's disease; epilepsy;
 KW Parkinson's disease; autoimmune disorder; inflammatory disorder; allergy;
 KW Crohn's disease; transgenic animal; animal model.
 XX

OS Homo sapiens.
 XX
 XX WO2003000864-A2.
 XX
 PD 03-JAN-2003.
 XX
 XX 20-JUN-2002; 2002WO-US021179.
 XX
 XX 22-JUN-2001; 2001US-0300518P.
 PR 29-JUN-2001; 2001US-0301787P.
 PR 29-JUN-2001; 2001US-0301792P.
 PR 29-JUN-2001; 2001US-0301892P.
 PR 29-JUN-2001; 2001US-0301893P.
 PR 06-JUL-2001; 2001US-0303405P.
 PR 06-JUL-2001; 2001US-0303442P.
 PR 15-MAR-2002; 2002US-0364438P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Gandhi AR, Swarnakar A, Hafalia AJA, Warren BA, Emerling BM;
 PI Arvizu CS, Ison CH, Honchell CD, Lee EA, Yue H, Forsythe IJ;
 PI Ramkumar J, Griffin JA, Yang J, Sanjanwala MM, Baughn MR;
 PI Borowsky ML, Yao MG, Walia NK, Bandman O, Lal PG, Becha SD, Lee SY;
 PI Richardson TW, Elliott VS, Luo W, Tang YT, Zebardjian Y, Lu Y;
 XX
 XX WPI; 2003-201420/19.
 DR N-PSDB; AAL51588.
 XX
 XX New nucleic acid-associated proteins and polynucleotides, useful for
 PT diagnosing, treating or preventing cell proliferative (e.g. cancer),
 PT neurological (e.g. epilepsy or Parkinson's disease), or autoimmune
 PT disorders (e.g. AIDS).
 XX
 XX Claim 1; Page 273-274; 312pp; English.
 XX
 XX The invention comprises the amino acid and coding sequences of human
 CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
 CC the invention are useful for diagnosing, treating or preventing disorders
 CC associated with aberrant expression of NAAP, such as: cell proliferative
 CC disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis
 CC or cancer); developmental disorders (e.g. renal tubular acidosis, anaemia
 CC or mental retardation); neurological disorders (e.g. Alzheimer's disease,
 CC Parkinson's disease or epilepsy); and autoimmune/inflammatory disorders
 CC (e.g. AIDS, allergies, asthma or Crohn's disease). The DNA sequences of
 CC the invention are useful for creating transgenic animals to model human
 CC disease. The present amino acid sequence represents a human nucleic acid-
 CC associated protein of the invention
 XX
 XX Sequence 480 AA;
 SQ

Query Match 91.6%; Score 2556; DB 6; Length 480;
 Best Local Similarity 92.7%; Pred. No. 2e-174;
 Matches 480; Conservative 0; Mismatches 0; Indels 38; Gaps 1;

QY 1 MGSPAPEGALGYVRETRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACS 60
 DB 1 MGSPAPEGALGYVRETRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACS 60
 QY 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYQLM 120
 DB 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYQLM 120
 QY 121 EHVVOACHRFIOASVEPLGISLRPLEAEPTPTAPPGRSRRSEGHDPPTESRSCSQ 180
 DB 121 EHVVOACHRFIOASVEPLGISLRPLEAEPTPTAPPGRSRRSEGHDPPTESRSCSQ 180
 QY 181 PPSASPDPKACNWKYKIVLNSQAQSLVGRSSGQPCQARLPSCGDEASSSSSS 240
 DB 181 PPSASPDPKACNWKYKIVLNSQAQSLVGRSSGQPCQARLPSCGDEASSSSSS 240
 QY 241 SSSSSEEGPIPGQSRISPTAATVQKCGAPASTPYLLTQAQDTSGSPSERARPLPGSE 300
 DB 241 SSSSSEEGPIPGQSRISPTAATVQKCGAPASTPYLLTQAQDTSGSPSERARPLPGSE 300

QY 301 FFSQNCCEAVAGCSGLDLSLVPGDEDPKPKQCLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 Db 301 FFSQNCCEAVAGCSGLDLSLVPGDEDPKPKQCLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 QY 361 ICGARFNRPNALKTSHRIHSGEKPKYKCTCGSRFVQVRSQPPSGFQKPARGGVGQKGGF 420
 Db 361 ICGARFNRPNALKTSHRIHSGEKPKYKCTCGSRFVQVRSQPPSGFQKPARGGVGQKGGF 420
 QY 421 CSSQRLKSPPSOVAHLRAHVLHTGKPKYPCPTCTGTRPRHLQTLKSHVRIHTGKPYH 480
 Db 396 -----QVAHLRAHVLHTGKPKYPCPTCTGTRPRHLQTLKSHVRIHTGKPYH 442
 QY 481 CDPCLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
 Db 443 CDPCLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 480

RESULT 3
 ABU03467
 ID ABU03467 standard; protein; 479 AA.
 XX AC ABU03467;
 XX DT 21-JAN-2003 (first entry)
 XX XX Angiogenesis-associated human protein sequence #12.
 DE Human; angiogenesis-associated transcript; angiogenesis;
 KW angiogenesis-associated disease; cancer; cytostatic.
 KW Homo sapiens.
 OS
 XX WO200279492-A2.
 XX PD 10-OCT-2002.
 XX
 XX 14-FEB-2002; 2002WO-US004915.
 XX PR 14-FEB-2001; 2001US-00784356.
 XX PR 22-FEB-2001; 2001US-00791390.
 XX PR 19-APR-2001; 2001US-0285475P.
 XX PR 03-AUG-2001; 2001US-0310025P.
 XX PR 13-NOV-2001; 2001US-0350666P.
 XX PR 29-NOV-2001; 2001US-0334244P.
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Murray R, Glynn R, Watson SR, Aziz N;
 XX WPI; 2003-040681/03.
 XX N-PSDB; ABX08750.
 XX
 XX Detecting angiogenesis-associated transcript in a cell for diagnosing and
 PT treating cancer by contacting a sample with a polynucleotide that
 PT exhibits changes in expression level as a function of time in tissue
 PT undergoing angiogenesis.
 XX
 XX Example 2; Page 190; 291pp; English.
 XX
 XX The present invention relates to methods and compositions for detecting
 CC an angiogenesis-associated transcript in a cell in a patient. The method
 CC involves contacting a biological sample from the patient with a
 CC polynucleotide that selectively hybridises to a sequence at least 80%
 CC identical to any of the angiogenesis-associated human polynucleotide
 CC sequences given in the specification. These angiogenesis-associated
 CC polynucleotide sequences comprise genes that exhibit changes in
 CC expression levels as a function of time in tissue undergoing
 CC angiogenesis. The method and the polynucleotide sequences of the
 CC invention are useful for diagnosing and treating angiogenesis and
 CC angiogenesis-associated diseases e.g. cancer. The polynucleotide
 CC sequences are also useful in the gene therapy of such disorders. The
 CC angiogenesis-associated proteins encoded by the polynucleotide sequences

CC are useful as a vaccine for therapeutic and prophylactic immunisation.
 CC ABU03456-ABU03569 represent angiogenesis-associated protein sequences
 XX
 SQ Sequence 479 AA;
 Query Match 91.1%; Score 2541.5; DB 6; Length 479;
 Best Local Similarity 92.5%; Pred. No. 2.2e-173;
 Matches 479; Conservative 0; Mismatches 0; Indels 39; Gaps 2;
 QY 1 MGSFAPEAGLGVVRFTHSSDVLGNLNLRLGILTDVTLVGGOPLAHKAVALTACS 60
 Db 1 MGSFAPEAGLGVVRFTHSSDVLGNLNLRLGILTDVTLVGGOPLAHKAVALTACS 60
 QY 61 GFYSIFRGRAGVGVVLSLPGGPEARGFAPLDDFMYTSRLRLSPATAPAVLAATYLOM 120
 Db 61 GFYSIFRGRAGVGVVLSLPGGPEARGFAPLDDFMYTSRLRLSPATAPAVLAATYLOM 120
 QY 121 EHVQACHRFIOASYEPLGISLRPLEAEPTTAPPGSPRSEGHDPPTESRSCSQ 180
 Db 121 EHVQACHRFIOASYEPLGISLRPLEAEPTTAPPGSPRSEGHDPPTESRSCSQ 180
 QY 181 PPSASPDPKACNWKYKIVLNSQASQAGSLVGRSSGQPCQARLPDGDEASSSSSS 240
 Db 181 PPSASPDPKACNWKYKIVLNSQASQAGSLVGRSSGQPCQARLPDGDEASSSSSS 239
 QY 241 SSSSSEEGIPGQSRSLPTAATVQKCGAPASTPYLLTSAQDTSGSPERARPLPGSE 300
 Db 240 SSSSSEEGIPGQSRSLPTAATVQKCGAPASTPYLLTSAQDTSGSPERARPLPGSE 299
 QY 301 FFSQNCCEAVAGCSGLDLSLVPGDEDPKPKQCLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 Db 300 FFSQNCCEAVAGCSGLDLSLVPGDEDPKPKQCLCRSSFRYKGNLASHRTVHTGKPYHCS 359
 QY 361 ICGARENRPANLKTSHRIHSGEKPKYKCTCGSRFVQVRSQPPSGFQKPARGGVGQKGGF 420
 Db 360 ICGARENRPANLKTSHRIHSGEKPKYKCTCGSRFVQVRSQPPSGFQKPARGGVGQKGGF 394
 QY 421 CSSQRLKSPPSOVAHLRAHVLHTGKPKYPCPTCTGTRPRHLQTLKSHVRIHTGKPYH 480
 Db 395 -----QVAHLRAHVLHTGKPKYPCPTCTGTRPRHLQTLKSHVRIHTGKPYH 441
 QY 481 CDPCLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
 Db 442 CDPCLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 479

RESULT 4
 ABU64323
 ID ABU64323 standard; protein; 479 AA.
 XX AC ABU64323;
 XX DT 11-MAR-2004 (first entry)
 XX DE AAC2-2 protein.
 XX KW Angiogenesis; cancer; AAC2-1; AAC2-2; tumour antigen; expression vector;
 KW cytostatic; gene therapy.
 XX OS Unidentified.
 XX PN WO2003080800-A2.
 XX PD 02-OCT-2003.
 XX PF 20-MAR-2003; 2003WO-US008536.
 XX PR 20-MAR-2002; 2002US-0365982P.
 XX PA (AVET) AVENTIS PASTEUR INC.
 XX PI Berinstein N, Lovitt C, Farrington M, Pedyczak A, Radvanyi L;
 PI Singh-Sandhu D;

Db 1 MGSPAPEGALGVYREFTRHSSDVLGNLNEIRLGIILTDVLLVGGQPLRAHKAVALIACS 60
 QY 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYLOM 120
 Db 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYLOM 120
 QY 121 EHVQACHRFIOASYEPLGISLRPLEAEPPTTAPPPGSPRRSEGHDPPTESRSCSQ 180
 Db 121 EHVQACHRFIOASYEPLGISLRPLEAEPPTTAPPPGSPRRSEGHDPPTESRSCSQ 180
 QY 181 PPSASPDPKACNMWKYKIVLNSQASGLVGRSSGQPCQARLPSPGDEASSSSSS 240
 Db 181 PPSASPDPKACNMWKYKIVLNSQASGLVGRSSGQPCQARLPSPGDEA-SSSSSS 239
 QY 241 SSSSEEGPIPGQSRSLPTAATVQFKCGAPASTPYLLTSAQDTSGPSERARPLPGE 300
 Db 240 SSSSEEGPIPGQSRSLPTAATVQFKCGAPASTPYLLTSAQDTSGPSERARPLPGE 299
 QY 301 FFSQNCCEAVAGCSGLDLSVPGDEDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 Db 300 FFSQNCCEAVAGCSGLDLSVPGDEDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 359
 QY 361 ICGARFNRPANLKTSHRIHSGEKPYKCTCGSRFVQVRSQPPSGFQKPARGGVGQKGF 420
 Db 360 ICGARFNRPANLKTSHRIHSGEKPYKCTCGSRFV- 394
 QY 421 CSSQRODLKSPSQVAHLRAHVLHTGKPYKPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
 Db 395 -----QVAHLRAHVLHTGKPYKPCPTCGTRFRHLQTLKSHVRIHTGKPYH 441
 QY 481 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
 Db 442 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 479

RESULT 6

ABU64322
 ID ABU64322 standard; protein; 480 AA.
 XX
 AC ABU64322;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE AAC2-1 protein.
 XX
 KW Angiogenesis; cancer; AAC2-1; AAC2-2; tumour antigen; expression vector;
 KW Cytostatic; gene therapy.
 XX
 OS Unidentified.
 XX
 FN WO2003080800-A2.
 XX
 PD 02-OCT-2003.
 XX
 PF 20-MAR-2003; 2003WO-US008536.
 XX
 PR 20-MAR-2002; 2002US-0365982P.
 XX
 PA (AVET) AVENTIS PASTEUR INC.
 XX
 PI Berinstein N., Lovitt C., Parrington M., Pedyczak A., Radvanyi L;
 PI Singh-Sandhu D;
 XX
 DR WPI; 2003-779251/73.
 XX
 DR N-PSDB; AAL56276.
 XX
 PT New expression vector for preventing or treating an angiogenesis-
 PT dependent disease (e.g. breast cancer) comprises a nucleic acid sequence
 PT that encodes the angiogenesis-associated antigen AAC2-1 or AAC2-2.
 XX
 PS Claim 50; Page 56; 57pp; English.
 XX

CC The present invention relates to an expression vector for inducing an
 CC anti-tumour immune response in a patient. Also provided is a method of
 CC using such a vector to treat angiogenesis-dependent diseases. The
 CC composition and methods are useful in diagnosing, preventing, prognosing
 CC or treating an angiogenesis-dependent disease, preferably a breast
 CC cancer. The DNA molecule and protein may also be used in drug screening
 CC assays. The present sequence is an AAC2 protein sequence shown in the
 CC exemplification of the invention
 XX
 SQ Sequence 480 AA;

Query Match 88.3%; Score 2464; DB 7; Length 480;
 Best Local Similarity 89.8%; Pred. No. 7.6e-168;
 Matches 465; Conservative 2; Mismatches 13; Indels 38; Gaps 1;

QY 1 MGSPAPEGALGVYREFTRHSSDVLGNLNEIRLGIILTDVLLVGGQPLRAHKAVALIACS 60
 Db 1 MGSPAPEGALGVYREFTRHSSDVLGNLNEIRLGIILTDVLLVGGQPLRAHKAVALIACS 60
 QY 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYLOM 120
 Db 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATYLOM 120
 QY 121 EHVQACHRFIOASYEPLGISLRPLEAEPPTTAPPPGSPRRSEGHDPPTESRSCSQ 180
 Db 121 EHVQACHRFIOASYEPLGISLRPLEAEPPTTAPPPGSPRRSEGHDPPTESRSCSQ 180
 QY 181 PPSASPDPKACNMWKYKIVLNSQASGLVGRSSGQPCQARLPSPGDEASSSSSS 240
 Db 181 PPSASPDPKACNMWKYKIVLNSQASGLVGRSSGQPCQARLPSPGDEASSSSSS 240
 QY 241 SSSSEEGPIPGQSRSLPTAATVQFKCGAPASTPYLLTSAQDTSGPSERARPLPGE 300
 Db 241 SSSSEEGPIPGQSRSLPTAATVQFKCGAPASTPYLLTSAQDTSGPSERARPLPGE 300
 QY 301 FFSQNCCEAVAGCSGLDLSVPGDEDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 Db 301 FSAARTVRLWQGAHRLDLSVPGDEDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 QY 361 ICGARFNRPANLKTSHRIHSGEKPYKCTCGSRFVQVRSQPPSGFQKPARGGVGQKGF 420
 Db 361 ICGARFNRPANLKTSHRIHSGEKPYKCTCGSRFV- 395
 QY 421 CSSQRODLKSPSQVAHLRAHVLHTGKPYKPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
 Db 396 -----QVAHLRAHVLHTGKPYKPCPTCGTRFRHLQTLKSHVRIHTGKPYH 442
 QY 481 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
 Db 443 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 480

RESULT 7

ABM84710
 ID ABM84710 standard; protein; 458 AA.
 XX
 AC ABM84710;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human diagnostic and therapeutic pproteins
 XX
 KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 XX
 OS Homo sapiens.
 XX
 FN WO2004023973-A2.
 XX
 PD 25-MAR-2004.
 XX
 PF 12-SEP-2003; 2003WO-US028227.
 XX
 PS 12-SEP-2002; 2002US-0410259P.

PR 12-SEP-2002; 2002US-0410260P.
 XX (INCY-) INCYTE CORP.
 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Harthshorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;
 PI Mooney EM, Deleane AM, Paneser IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patury S, Shi X, Suarez C;
 XX WPI; 2004-329368/30.
 DR N-PSDB; ACN43362.
 DR
 XX
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 XX Claim 27; Page; 190pp; English.
 PS
 XX The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 XX Sequence 458 AA;
 XX
 Query Match 86.6%; Score 2414; DB 8; Length 458;
 Best Local Similarity 88.4%; Pred. No. 2.7e-164;
 Matches 458; Conservative 0; Mismatches 0; Indels 60; Gaps 3;
 XX
 QY 1 MGSPAPEGALGYVREFTTHSSDVLGNLNLRLGILTDVTLVGGQPLRAHKAVALIACS 60
 DB 1 MGSPAPEGALGYVREFTTHSSDVLGNLNLRLGILTDVTLVGGQPLRAHKAVALIACS 60
 QY 61 GFYISIFGRAGVGVVDVLSLPGGPEARGFAPLLDFMYTSRLSPATAPAVLAATYILQM 120
 DB 61 GFYISIFGRAGVGVVDVLSLPGGPEARGFAPLLDFMYTSRLSPATAPAVLAATYILQM 120
 QY 121 EHVQACHRFIQASYEPLGISLRLEAEPTPTAPPGSPRRSGHPDPTESRCSQG 180
 DB 121 EHVQACHRFIQASYEPLGISLRLEAEPTPTAPPGSPRRSGHPDPTESRCSQG 180
 QY 181 PPSAPDPKACNWKYKIVLNSQASQAGSLVGRSSQCPQARLSGDEASSSSSS 240
 DB 181 PPSAPDPKACNWKYKIVLNSQASQAGSLVGRSSQCPQARLSGDEASSSSSS 218
 QY 241 SSSSEEGPIPGQSRSLPTAATVQKCAPASTPYLLTSQAQDTSGSPSERARPLGSE 300
 DB 219 SSSSEEGPIPGQSRSLPTAATVQKCAPASTPYLLTSQAQDTSGSPSERARPLGSE 278
 QY 301 PFSCQCEAVAGCSGLDVLPGDEBDKPYKCOLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 DB 279 PFSCQCEAVAGCSGLDVLPGDEBDKPYKCOLCRSSFRYKGNLASHRTVHTGKPYHCS 338
 QY 361 ICGARNRPNLKTSHRISHGKPKYKCTCGSRFVQVRSQPPSGFGKARGGVQKGGF 420
 DB 339 ICGARNRPNLKTSHRISHGKPKYKCTCGSRFV----- 373

QY 421 CSSQODLKSPSPQVAHLRAHLVLIHTGKPKYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
 DB 374 -----QVAHLRAHLVLIHTGKPKYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 420
 QY 481 CDPGLHFRHKSQRLHLRLQKHGAATNTKVHYHILGGP 518
 DB 421 CDPGLHFRHKSQRLHLRLQKHGAATNTKVHYHILGGP 458
 RESULT 8
 ABB57289
 ID ABB57289 standard; protein; 707 AA.
 XX
 AC ABB57289;
 XX
 DT 07-MAR-2002 (first entry)
 XX
 DE Mouse ischaemic condition related protein sequence SEQ ID NO:814.
 XX
 KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 XX vasospastic ischaemia; ischaemic condition; ischaemic disease.
 OS Mus musculus.
 XX
 PN WO200188188-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-JP004192.
 XX
 PR 18-MAY-2000; 2000JP-00145977.
 XX
 PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX WPI; 2002-034733/04.
 DR N-PSDB; ABI99745.
 XX
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 PT genes.
 XX
 PS Claim 2; Page 2004-2007; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring the
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used
 CC in the exemplification of the present invention
 XX
 SQ Sequence 707 AA;
 Query Match 40.8%; Score 1138; DB 5; Length 707;
 Best Local Similarity 36.4%; Pred. No. 9.1e-73;
 Matches 269; Conservative 51; Mismatches 149; Indels 270; Gaps 14;
 QY 1 MGSPAPEGALGYVREFTTHSSDVLGNLNLRLGILTDVTLVGGQPLRAHKAVALIACS 60
 DB 1 MASPA--DSCI-----QFTRHASDVLNLRSLRSDILTDVVIVVSREQFRAHKTVMACS 54
 QY 61 GFYISIFGRAGVGVVDVLSLPGGPEARGFAPLLDFMYTSRLSPATAPAVLAATYILQM 120

Db 55 GLFYISFTDQKCNLSVINLDPSEISPEGFCILLDFMYTSRLNLRGNIMAVMTTAMYLQM 114
Qy 121 EHVQACHRFICASVEPLGISRLPLEAE----- 148
Db 115 EHVVDTCRKFIKASEAMAPALKPREEFLNSRMLMHPDIMAYRGVVENNNPLRPTPG 174
Qy 149 -----PPT----- 151
Db 175 CESRAPAPLYSGLSTPPASYPWYSHLPLSTFLSDEELRDAPRMPVANPPKRALPCD 234
Qy 152 ----- 151
Db 235 SARQVNEYSRAMEVPSLCHSNISYKPAEVPPEARSDIHYSPGPKPAVPSARNAPY 294
Qy 152 -----PPTAP-----PGSPRRSGHPDPPTESRS----- 176
Db 295 PFCDKASKEERPSSEDEIHALFEPNAPLNKGLVSPQSPKSDCQPNPSPTSCSSKNA 354
Qy 177 -----CSQPPSPASPPDKACNWKYKIVINS-----QASQAGSLVGRSSGQPCQA 225
Db 355 CILQASGSPPAKSPPTDKACNWKYKIVINSLNQNAKPEGSEQAELGRLSPRAYPAPPA 414
Qy 226 RLPSGDEAS--SSSSSSSSSEEGIPGQSRLL-----SPTAATVQF----- 266
Db 415 CQPPMEPANLDLSQPTKLSASGEDSTIP-QASRLNANLNRSLGSGPSRSSSESHSPLYMHP 473
Qy 267 -KCGAPASTPYLLTQAQDTSQSPSERAPLPQSEF-----PSCONCEAVAGCSSG 316
Db 474 PKTSCGSGSQPQTEMLTAGFTPEEMGETQSEYSSSCSNGTFFCNECDRFSSEAS 533
Qy 317 LD-SLVPGDEDPYKQLCRSPFRYKGNLASHRTVHTGKPYHCISGAFNRPNANLKH 375
Db 534 LKHILQTHSDPYKCDRCQASFRYKGNLASHRTVHTGKPYHCISGAFNRPNANLKH 593
Qy 376 SRIHSGEPKPYCETCGSRFVQVRSPGPGFQKARGGVQKGGFCSSQORDLKSPPPQV 435
Db 594 TRIHSGEPKPYCETCGARFV-----QV 615
Qy 436 AHLRAVLHTGKPYCETCTRFRHLQTLKSHVRIHTGKPYHCDCGLHFRKHSQLR 495
Db 616 AHLRAVLHTGKPYCETCTRFRHLQTLKSHVRIHTGKPYHCDCGLHFRKHSQLR 675
Qy 496 LHLRKHGAATWKVHYHI 514
Db 676 LHLRKHGAATWKVQYRV 694

RESULT 9
ADRI14065
XX ADRI14065 standard; protein; 706 AA.
AC ADRI14065;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human NF-kappaB pathway-associated protein SeqID66.
XX
KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnary; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;

KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; human.
OS Homo sapiens.
XX WO20004065577-A2.
XX
XX 05-AUG-2004.
XX
XX 13-JAN-2004; 2004WO-US000798.
XX
XX 14-JAN-2003; 2003US-0440068P.
XX 12-MAY-2003; 2003US-0469577P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Nadler SG, Neubauer MG, Feder JN, Carman J;
XX N-PSDB; ADRI14064.
XX
XX WPI; 2004-562168/54.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.
XX
XX Claim 6; SEQ ID NO 66; 237pp; English.
XX
XX This invention relates to the novel association of protein sequences (and
XX the genes which encode them) to the NF-kappaB pathway. The invention may
XX be useful for the production of compounds with an antiinflammatory,
XX cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
XX gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
XX vulnary activity or for gene therapy. The proteins and nucleotides are
XX useful for diagnosing, preventing, treating, or ameliorating conditions
XX or diseases associated with the NF-kappaB pathway. The condition is an
XX immune disorder, an inflammatory disorder, an inflammatory disorder
XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
XX hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory
XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
XX immune activity, disorders related to aberrant acute phase responses,
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,
XX organ transplant rejection, conditions related to organ transplant
XX rejection, disorders related to aberrant signal transduction,
XX proliferating disorders, cancers and HIV propagation in cells infected
XX with other viruses. The present sequence is that of a human protein which
XX is subject to the novel association with the NF-kappaB pathway of the
XX invention. Note: This sequence does not appear in the specification but
XX was obtained by the indexer from Genbank.
XX
XX SQ Sequence 706 AA;

Query Match 40.5%; Score 1128.5; DB 8; Length 706;
Best Local Similarity 35.0%; Pred. No. 4.4e-72;
Matches 270; Conservative 48; Mismatches 117; Indels 337; Gaps 16;
Qy 1 MGSPAAPGALGYVREFTRHSSDVLGNLNEURLRGLITDVTLLVGGQPLRAHKAVLIACS 60
Db 1 MASPA--DSCI---QFTRHSDVLLNLRSLRDLTDVVIVVSREQFRAHKTVLMACS 54
Qy 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLDFMYTSRLNLRSPATPAVLAATYLOM 120
Db 55 GLFYISFTDQKCNLSVINLDPSEISPEGFCILLDFMYTSRLNLRGNIMAVMTTAMYLQM 114
Qy 121 EHVQACHRFICASVEPLGISRLPLEAE----- 148
Db 115 EHVVDTCRKFIKASEAMAPALKPREEFLNSRMLMHPDIMAYRGVVENNNPLRPTPG 174

QY 135 -----YEPGIS----- 141
 Db 175 CESRAFAPSLYGLSTPPASYSMYSHLPVSLFLSDEFRDVRMPVANPFPKERALPCDS 234
 QY 142 -----LRPL----- 145
 Db 235 ARPVGYSRPTLEVPNVCHSNISPKETIPEARSDMHYSVAEGLKPAAPSARNAPYF 294
 QY 146 -----EAPPT-----PPTAP-----PPGSRSEGHDPPTESRS----- 176
 Db 295 PCDKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSKDCQPNPTSCSKNAC 354
 QY 177 ---CSOGPPSPAPDPKACNWKYKIVLNSQASQAGSLVGRSSQCPQARLPSGDEA 233
 Db 355 ILQASGSPPAKSPDTPKACNWKYKIVLN----- 384
 QY 234 SSSSSSSSSSEBGPFGQSRSLPTAATVQFKCGAP-----ASTPYLLTSAQADT- 285
 Db 385 -----SLNQNAKEGPEQAEGLRLSPRAYTAPPACQPMPEPNLDLQSTKLASGEDST 439
 QY 286 -----SGSP-----SERARPL----- 296
 Db 440 IPQASRLNNIVNRMTGSPRSSSESHSLYMHPPKCTSCGSPQHAEMCLHTAGTPTPE 499
 QY 297 -----PGSEF-----FSCQCEAVAGCSSGLD-SLVPGEDEKPKYKQCRSFRYKG 342
 Db 500 EMGETOSEYSDSCENGAFECNEDCRFEEASLKRHTLQTHSDKPKYKCDRCQASFRYKG 559
 QY 343 NLASHRTVHTGKPYHCSICGAFENRPNANKTHSRHSKEKPKYKCTGSRFVQVRSQPP 402
 Db 560 NLASHRTVHTGKPYKNCICGAFENRPNANKTHSRHSKEKPKYKCTGSRFV----- 612
 QY 403 SGFQKGARGGVGQKGFCSQSORDLKPPSQVAHLRAHVLHTGKPYKPCPTGTRFRH 462
 Db 613 -----QVAHLRAHVLHTGKPYKPCPTGTRFRH 641
 QY 463 LOTLKSHTVHTGKPYKPCPTGTRFRHKSQRLHLRQKHGAATWKVYHI 514
 Db 642 LOTLKSHTVHTGKPYKPCPTGTRFRHKSQRLHLRQKHGAATWKVYRV 693
 RESULT 10
 ADJ70182
 ID ADJ70182 standard; protein; 706 AA.
 AC ADJ70182;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SeqID1988.
 XX
 KW mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX
 OS Homo sapiens.
 XX
 XX WO2003087768-A2.
 XX
 XX 23-OCT-2003.
 PD
 XX 04-APR-2003; 2003WO-US010870.
 XX
 XX 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-038987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 XX (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 XX WPI; 2003-845369/78.
 DR
 XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 with the disease.
 XX
 PS Claim 1; SEQ ID NO 1988; 180pp; English.
 XX
 CC This invention relates to novel mitochondrial targets that can be used
 for therapeutic intervention in treating a disease associated with
 altered mitochondrial function. Specifically, it refers to a method for
 identifying proteins of the human heart mitochondrial proteome that are
 useful for drug screening assays, as well as therapeutic targets. The
 present invention describes a method for identifying such proteins that
 can be used in the treatment of various diseases associated with altered
 mitochondrial function including diabetes mellitus, Huntington's disease,
 osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 compositions have neuroprotective, nootropic, antidiabetic,
 anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 cytostatic activities. This polypeptide sequence is a human heart
 mitochondrial protein of the invention.
 XX
 SQ Sequence 706 AA;
 Query Match 40.4%; Score 1126.5; DB 7; Length 706;
 Best Local Similarity 35.0%; Pred. No. 6.1e-72;
 Matches 270; Conservative 48; Mismatches 117; Indels 337; Gaps 16;
 QY 1 MGSPAAEGALGYVRETRHSSDVLGNELRLGILTDVLLVGGQPLRAHKAVALIACS 60
 Db 1 MASPA--DSCI---QFTRHASDVLNLRSLRSDILTDVVIVVSREQFRAHKTVMACS 54
 QY 61 GFFYSIFGRAGVGVVLSLPGGPEARFAPLLDFMTSLRLSPATAPAVLAATYQLM 120
 Db 55 GLFYSIFTDQLKCNLSVINLDPEINPEGFCILLDFMTSLRLNREGNIMAVATYQLM 114
 QY 121 EHVVOACHRFIQAS----- 134
 Db 115 EHVVDTCRKFKIKASEAMVSAIKPPREBFLNSRMLPQDIMAYRGREVNNPLRSAPG 174
 QY 135 -----YEPGIS----- 141
 Db 175 CESRAFAPSLYGLSTPPASYSMYSHLPVSLFLSDEFRDVRMPVANPFPKERALPCDS 234
 QY 142 -----LRPL----- 145
 Db 235 ARPVGYSRPTLEVPNVCHSNISPKETIPEARSDMHYSVAEGLKPAAPSARNAPYF 294
 QY 146 -----EAPPT-----PPTAP-----PPGSRSEGHDPPTESRS----- 176
 Db 295 PCDKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSKDCQPNPTSCSKNAC 354
 QY 177 ---CSOGPPSPAPDPKACNWKYKIVLNSQASQAGSLVGRSSQCPQARLPSGDEA 233
 Db 355 ILQASGSPPAKSPDTPKACNWKYKIVLN----- 384
 QY 234 SSSSSSSSSSEBGPFGQSRSLPTAATVQFKCGAP-----ASTPYLLTSAQADT- 285
 Db 385 -----SLNQNAKEGPEQAEGLRLSPRAYTAPPACQPMPEPNLDLQSTKLASGEDST 439
 QY 286 -----SGSP-----SERARPL----- 296
 Db 440 IPQASRLNNIVNRMTGSPRSSSESHSLYMHPPKCTSCGSPQHAEMCLHTAGTPTPE 499
 QY 297 -----PGSEF-----FSCQCEAVAGCSSGLD-SLVPGEDEKPKYKQCRSFRYKG 342

PF 15-OCT-1999; 99US-00418640.
 PR 15-OCT-1999; 99US-00418640.
 PA (ISIS-) ISIS PHARM INC.
 XX Taylor JK, Cowsett LM;
 XX WPI; 2001-048959/06.
 DR N-PSDB; AAC81137.
 XX
 PT Antisense compounds which specifically hybridize with and inhibit human
 PT bcl-6 expression, useful for treating bcl-6 related disorders, and
 PT preventing or delaying inflammation or tumor formation.
 PS Disclosure; Col 47-52; 42pp; English.
 XX
 CC This sequence represents human bcl-6. Bcl-6 (also known as B-cell
 CC CLL/lymphoma 6, zinc finger protein 51 and laz3) is a sequence- specific
 CC DNA-binding transcriptional repressor. The bcl-6 gene is expressed in
 CC germinal centre B- and T- cells and is required for germinal centre
 CC formation and Th-2 mediated antibody affinity maturation. Bcl-6 may also
 CC play a role in the regulation of apoptosis. The bcl-6 gene is located on
 CC chromosome 3q27, a region which undergoes a high frequency of
 CC translocation events. Such chromosomal translocations can result in
 CC aberrant forms of bcl-6, which are strongly implicated in the
 CC pathogenesis of several types of lymphoma, and have also been reported in
 CC acute lymphoblastic leukaemia and post-transplant lymphoproliferative
 CC disorders. The invention relates to antisense oligonucleotides targeted
 CC to the human bcl-6 gene, which inhibit its expression. A series of
 CC oligonucleotides (AAC81144-C81223) were designed to target different
 CC regions of the human bcl-6 mRNA, and were analysed for their effect on
 CC bcl-6 mRNA levels by quantitative real-time PCR. The oligonucleotides of
 CC the invention are useful for diagnosis, prevention and treatment of
 CC conditions associated with aberrant forms of bcl-6, such as lymphomas,
 CC acute lymphoblastic leukaemia and post-transplant lymphoproliferative
 CC disorders
 XX
 SQ Sequence 706 AA;

Query Match 40.2%; Score 1122.5; DB 4; Length 706;
 Best Local Similarity 36.3%; Pred. No. 1.2e-71;
 Matches 268; Conservative 58; Mismatches 142; Indels 271; Gaps 17;

QY 1 MGSPAPEGAGVGVRETHSSDNLNLRGLTDTVLLVGQPLRAHKAVALIACS 60
 DB 1 MASPA--DSCI-----QFTHASDVLNLRSLRDLTDVIVSREQFRAKTVLMACS 54
 QY 61 GFYSIFGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPVLAATYLOM 120
 DB 55 GLFYSIFTDLKCNLSVINLDPENPEGFCILLDFMYTSRLNLRGNIMAVMATYLOM 114
 QY 121 EHVQACHRPQAS-----YPLGIS----- 134
 DB 115 EHVVDTCRKFIRKASEAMVSAIKPPREFINRMLMPQDIMAYRGREVVNNLPLRSAPG 174
 QY 135 -----YPLGIS----- 141
 DB 175 CESRAPAPSLYGLSTPPASYSWYSHLPVSSLLFSDERDVRVMPVAPFPKRALPCDS 234
 QY 142 -----LRPL----- 145
 DB 235 ARPVGYSRPTLEVPNVCHSNISPKETIPPEARSDMHYSVAEGLKPAAPSARNAPVF 294
 QY 146 -----EAEPT-----PPTAP-----PPGPRSEGHDPPTESRS----- 176
 DB 295 PCDKASKBERPSEDEIALHFEPNAPLNRLKGLVSPQSPQKSDQPNSPTEACSSKNAC 354
 QY 177 ---CSQGPSPASPPKACNWKYKIVLNS-----QASQAGSIVGRSSGQPCP 223
 DB 355 ILQASGSPSPPTDPKACNWKYKIVLNSLNQNAKPGGPQAGELSPRAYTAPPAC 414
 QY 224 QARLPBGDEASSSSSSSSSSSEEGPIPGQSL-----SPTAATVQF----- 266

DB 415 QPPM-EPENLDQSPTKLSASGEDTIP-QASRLNNIVNRSMTGSPRSSSESHSPLYMHP 472
 QY 267 ---KCG--APASTPYLLTSQ-----AQDTSSGSPERARPLPGSEFFSCQNEAVAGCCSG 316
 DB 473 PKTSCGSSQSPQHAEMCLHTAGPTFAEMGTSQSEYSSSCENGAFFNCCDCRESEAS 532
 QY 317 LD-SLVPGDEDEKPKYKQLCRSSFYKGNLASHRTVHTGEKPYHGSICGARNRPNALKTH 375
 DB 533 LKRHTLQTHSDKPKYKCDRCQASFRYKGNLASHRTVHTGEKPYRNCNCGAQNRPNALKTH 592
 QY 376 SRIHSGKPKYKCTCGSRFVVRSPQPSGFGOKPARGVGQKGFCCSSQORDLKSPPPSQV 435
 DB 593 TRIHSGKPKYKCTCGARFV-----QV 614
 QY 436 AHLRAHVLHTGKPYCPTCGTRFRLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQLR 495
 DB 615 AHLRAHVLHTGKPYCEICGTRFRLQTLKSHLRIHTGKPYHCKCNLHFRHKSQLR 674
 QY 496 LHLRQKGAATNTKVHYHI 514
 DB 675 LHLRQKGAITNTKVQYRV 693
 RESULT 13
 ID ADL82847 standard; protein; 706 AA.
 XX ADL82847;
 AC ADL82847;
 DT 17-JUN-2004 (first entry)
 XX Human PRO26296, SEQ ID 49.
 DE
 XX Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic;
 KW Antiallergic; Muscular; Neuroprotective; Nephrotropic; Antiinflammatory;
 KW Gene therapy; PRO; B cell related disorder; cancer;
 KW immune-mediated inflammatory disease; human.
 XX Homo sapiens.
 OS
 XX WO2004024097-A2.
 XX 25-MAR-2004.
 XX 15-SEP-2003; 2003WO-US029097.
 XX 16-SEP-2002; 2002US-0411392P.
 XX (GETH) GENENTECH INC.
 PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood WI;
 PI Wu TD;
 DR WPI; 2004-329389/30.
 DR N-PSDB; ADL82846.
 XX
 PT New PRO polypeptide, useful for diagnosing and treating a B cell related
 PT disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune
 PT mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.
 XX
 PS Claim 10; Fig 49; 695pp; English.
 CC The present invention relates to PRO proteins and their coding sequences.
 CC The PRO proteins are useful for diagnosing and treating a B cell related
 CC disorder, e.g. X-linked infantile hypogammaglobulinemia, polysaccharide
 CC antigen unresponsiveness, selective IgA deficiency, selective IgM
 CC deficiency, selective deficiency of IgG subclasses, immunodeficiency with
 CC hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's
 CC lymphoma, intermediate lymphoma, follicular lymphoma, type II
 CC hypersensitivity, rheumatoid arthritis, autoimmune mediated haemolytic
 CC anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or
 CC ankylosing spondylitis. The PRO proteins are also useful for preparing a

CC medicament for treating a condition that is responsive to the PRO
 CC protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO
 CC coding sequences are useful as hybridization probes in chromosome and
 CC gene mapping, in preparing PRO proteins, or in generating transgenic
 CC animals or knockout animals, which in turn are useful in the development
 CC and screening of therapeutically useful reagents.
 XX
 SQ Sequence 706 AA;

Query Match 40.2%; Score 1122.5; DB 8; Length 706;
 Best Local Similarity 36.3%; Pred. No. 1.2e-71;
 Matches 268; Conservative 58; Mismatches 142; Indels 271; Gaps 17;

QY 1 MGSPAPEAGLVYREFTHSSDVLGNLRLRGLITDVLVGGQPLRAHKAVALIACS 60
 DB 1 MASPA--DSCI---QFTHSDVLLNLRSLRDLITDVLVVSREQFRAHKTVMACS 54
 QY 61 GFYSIFRGRAGVDVLSLPGGPARGAPALDPMYTSRLRLSPATPAVLAATYLOM 120
 DB 55 GLFYSIFTDLKCNLSVINLDPENPEGFCILLDFMYTSRLNLRGNIMAVMATYMLQM 114
 QY 121 EHVQACHRFIOAS----- 134
 DB 115 EHVDTCKRFKASEAMVSAIKPPREFLNSRMLMQDIMAYRGREVNNLPLRSAPG 174
 QY 135 -----YPLGTS----- 141
 DB 175 CESRAPAPSLYSGLSTPPASYSMYSHLPVSSLLFSDERFDRVMPVNPFPKRALPCDS 234
 QY 142 -----LRLP----- 145
 DB 235 ARPVGESYRPLEVSPNVCHSNISPKETIPPEARSDMHYSVAEGLKPAAPSARNAPYF 294
 QY 146 -----EAPPT-----PPTAP-----PPGSRSEGHDPPTESRS----- 176
 DB 295 PCDKASKEERPSSEDEIALHFEPPNAPLNRKGLVSPQSPQSDCQPNSTPEACSSKNAC 354
 QY 177 ---CSQGPSPASPPKACNWKYKYLNS-----QASQAGSLVGRSSGQPCP 223
 DB 355 ILQAGSGPPAKSPPTPKACNWKYKYLNSLNQNAKPGGPQEQASLGRSLSPRAYTAPAC 414
 QY 224 QARLPSGEASSSSSSSSSEEGPIPGQSRLL-----SPTAATVQF----- 266
 DB 415 OPPM-EPENLDLQSPKLSAGEDSTIP-QASRLANNVNRSMVTGSPRSSSHSPLYMHP 472
 QY 267 ----KCG--APASTVLLTSQ----AQDTSGSPSERARPLPGSEFPSCONCAVAGCSSG 316
 DB 473 PKCTSGSOPQAEMLHTAGTFAEEMGETQSEYSDSSCENGAFPCNECDRCFSEAS 532
 QY 317 LD-SLVPGDEDPYKCOLCRSSFRYKGNLASHRTVHTGKPYHCSICGARENRPANLAKTH 375
 DB 533 LKRHTLQTHSDKPYKCDRCQASFRYKGNLASHRTVHTGKPYRCNICGAQFNRPANLAKTH 592
 QY 376 SRIHSGEKPKCTCGSRVQVRSPPSPGFGKPARGGVQKGGFCSSQRDLKSPPSQV 435
 DB 593 TRIHSGEKPKCTCGARV-----QV 614
 QY 436 AHLRAVLHTGKPYPCPTCTRFRHLQTLKSHVRIHTGKPYKPCDGLHFRHKSQIR 495
 DB 615 AHLRAVLHTGKPYPCICGTRFRHLQTLKSHLRIHTGKPYHCEKCNLHFRHKSQIR 674
 QY 496 LHLRKHGAATWKVHYHI 514
 DB 675 LHLRKHGAITNTKVQYRV 693

RESULT 14
 ADRI4017
 ID ADRI4017 standard; protein; 706 AA.
 XX
 AC ADRI4017;
 XX
 DT 21-OCT-2004 (first entry)

XX Human NF-kappaB pathway-associated protein SeqID18.
 DE NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
 XX antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnery; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HIV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; human.
 XX Homo sapiens.
 OS
 XX WO2004065577-A2.
 XX 05-AUG-2004.
 PD 13-JAN-2004; 2004WO-US000798.
 PF 14-JAN-2003; 2003US-0440068P.
 PR 12-MAY-2003; 2003US-0469757P.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA Nadler SG, Neubauer MG, Feder JN, Carman J;
 PI WPI; 2004-562168/54.
 XX N-PSDB; ADR14016.
 DR New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders of
 PT diseases associated with NF-kappaB pathway.
 PS Claim 6; SEQ ID NO 18; 237pp; English.
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
 CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnery activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HIV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human protein which
 CC is subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

bcl-6 locus; non-Hodgkin lymphoma; B-cell lymphoma; B-lymphocyte; diagnostic; therapeutic; chromosome-3q27; translocation; proto-oncogene; diffuse large cell lymphoma; DLCL; zinc finger.

```

QY 146 -----EAPPT-----PPTAP-----PPGSPRRSEGHDPPTTES-----RSC 177
Db 295 PCDKASKEERESSEDEIHALHFEPPNAPLNKGLVSPQSPQKSDCQPNSTPCACSSKNAC 354
QY 178 -----SQGPPSPAPDPKACNWKYKIVLNS-----QASQAGSLVGERSSGQPCP 223
Db 355 ILQSGSPPAKSPDPKACNWKYKIVLNSLNONAKPGPEQABGLRLSPRAYTAPPAC 414
QY 224 QARLPSGDEASSSSSSSEEGPIPGQSR-----SPTAATVQF----- 266
Db 415 QPPM-EPENLDLQSPKUSAGEDSTIP-QASRLNINVRNMTGSPRSSSHSPLYMHP 472
QY 267 -----KCG--APASTPYLLTSQ-----AQDTSGSPSERARPLPGSEFFSQNCEAVAGCSSG 316
Db 473 PKCTSGSQSPQHAEWCLHTAGTFAESMGETQSEYSDSCENGAFFCNECDCRFSEAS 532
QY 317 LD-SLVPGDEDPYKQCLCRSSFRYKGNLASHRTVHTGKPYHCICGARFNRPANLKT 375
Db 533 LKRHTLQTHSDKPYKCDRCQASFRYKGNLASHKTVHTGKPYRCNICGAQFNRPANLKT 592
QY 376 SRIHSGKPYKCTCGSRPVQVRSQPPSGFQKPARGGVGQKGFCSQRQDLKSPPSOV 435
Db 593 TRIHSGKPYKCTCGARFV-----OV 614
QY 436 AHLRAHVLHTGKPYPCPTCGTRFRLHQLTKSHVRIHTGKPYHCDPCGLHFRHKSOLR 495
Db 615 AHLRAHVLHTGKPYPCPCICGTRFRLHQLTKSHLRIHTGKPYHCEKCNLHFRHKSOLR 674
QY 496 LHLRQKHGAATNTKVHYHI 514
Db 675 LHLRQKHGAITNTKVQYRV 693

```

Search completed: August 9, 2005, 14:02:13
Job time : 172 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 13:57:02 ; Search time 44 Seconds
(without alignments)
878.823 Million cell updates/sec

Title: US-09-815-379-4

Perfect score: 2789

Sequence: 1 MGSPAPEGALGVYRETRH.....RQKHGAATNTKVHYILGGP 518

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1128.5	40.5	706	4	US-09-538-092-1104
2	1095.5	39.3	706	1	US-08-074-967-2
3	1095.5	39.3	706	2	US-08-553-541B-2
4	1095.5	39.3	706	3	US-09-268-202-2
5	1095.5	39.3	706	4	US-09-761-117-2
6	1095.5	39.3	706	5	PCT-US94-06669-2
7	576.5	20.7	803	3	US-09-063-035-2
8	476.5	17.1	733	4	US-09-949-016-7651
9	457.5	16.4	641	4	US-09-949-016-10888
10	457.5	16.4	641	4	US-09-949-016-10889
11	457.5	16.4	641	4	US-09-949-016-10890
12	457.5	16.4	641	4	US-09-949-016-10891
13	428.5	15.4	678	4	US-09-949-016-10273
14	422.5	15.1	769	4	US-09-949-016-8149
15	421.5	15.1	458	4	US-09-538-092-1356
16	421.5	15.1	673	4	US-09-538-092-1270
17	419.5	15.0	488	3	US-09-933-750C-17
18	419.5	15.0	488	3	US-09-234-613-17
19	415	14.9	803	4	US-09-538-092-1026
20	412	14.8	639	4	US-09-949-016-9434
21	405.5	14.5	638	4	US-09-949-016-11391
22	401	14.4	1191	4	US-09-949-016-6356
23	399	14.3	469	4	US-09-538-092-1195
24	399	14.3	656	4	US-09-949-016-10075
25	396	14.2	685	4	US-09-949-016-8627
26	393	14.1	453	4	US-09-949-016-10252
27	392	14.1	292	4	US-09-538-092-1334

28	389	13.9	663	4	US-09-949-016-7484	Sequence 7484, Ap
29	388.5	13.9	698	4	US-09-949-016-11419	Sequence 11419, A
30	387	13.9	325	4	US-09-538-092-11386	Sequence 11386, Ap
31	386	13.8	654	4	US-09-538-092-1193	Sequence 1193, Ap
32	386	13.8	654	4	US-09-949-016-6357	Sequence 6357, Ap
33	385	13.8	185	4	US-09-494-190-125	Sequence 125, App
34	385	13.8	185	4	US-09-494-190-126	Sequence 126, App
35	384.5	13.8	547	1	US-08-340-203A-3	Sequence 3, Appli
36	384.5	13.8	547	2	US-08-452-567-3	Sequence 3, Appli
37	384.5	13.8	547	2	US-08-452-427-3	Sequence 3, Appli
38	384.5	13.8	547	3	US-09-085-407-3	Sequence 3, Appli
39	384.5	13.8	578	4	US-09-949-016-6715	Sequence 6715, Ap
40	383	13.7	583	4	US-09-949-016-9840	Sequence 9840, Ap
41	382.5	13.7	540	4	US-09-949-016-6358	Sequence 6358, Ap
42	382.5	13.7	564	4	US-09-949-016-11234	Sequence 11234, A
43	381.5	13.7	717	4	US-09-881-578A-2	Sequence 2, Appli
44	378.5	13.6	506	4	US-09-538-092-1177	Sequence 1177, Ap
45	375.5	13.5	671	3	US-09-121-321-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-538-092-1104
; Sequence 1104, Application US/09538092
; Patent No. 8753314
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 1104
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P41182
US-09-538-092-1104

Query Match 40.5%; Score 1128.5; DB 4; Length 706;

Best Local Similarity 35.0%; Pred. No. 1.8e-78;

Matches 270; Conservative 48; Mismatches 117; Indels 337; Gaps 16;

Qy	1	MGSPAPEGALGVYRETRHSSDVLGNLNLRLGILTDVLLVGGOPLRAHKAVLIACS	60
Db	1	MASPA-----QFTRHSDVLLNLRLSRDILTDVVIVSRQFRAKTVLACS	54
Qy	61	GFYISIFRAGVGVVLSLPGGPEARGFAPLDFMYTSRLRLSPATAPAVLAATYIQM	120
Db	55	GLFYSIFTDQLKNLSVINLDPEINPFGFCLLDPMYTSRLNLREGNIMAVMATAYIQM	114
Qy	121	EHVVOACHRIQAS-----	134
Db	115	EHVVDTCKRKIKASEAMVSAIKPPPEEFNLNRMPLQDINAYRGREVNNLPLRSAPG	174
Qy	135	-----YEPLGIS-----	141
Db	175	CESRAFAPSLYGLSTPPASYSMYSHLPVSSLLFDEDFRDVRMPVANPPFKERALPCDS	234
Qy	142	-----LRPL-----	145
Db	235	ARVPGEYSRPTLEVSPNVCHSNISPKETIPPEARSDMHVSAEGLKPAAPSARNADPYF	294

Db 175 CESRAFPALSYGLSTPSPASYSNMYSHLPVSSLLFSDEFRDVRMPVANPFPKRALPCDS 234
QY 142 -----LRLE----- 145
Db 235 ARPVPGEYSRPTLEVPNSVCHSNISYKETIPEARSMDHYSVAEGLKPAAPSARNAPYF 294
QY 146 -----EAEPT-----PPTAP-----PPGSRSEGHDPPTES-----RSC 177
Db 295 PCOKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSPKSDCQNSPTEACSSKNAC 354
QY 178 -----SQGPPSPASPPKACNWKYKIVLNS-----QASOAGSLVGRSSGQPCP 223
Db 355 ILQSGSPPAKSPPTDPKACSWKYKIVLNSLNQNAKPGPQAEGLRLSPRAYTAPPAC 414
QY 224 QARLPBGDEASSSSSSSSEEGIPQPSRL-----SPTAATVQF----- 266
Db 415 QPPM-EPENLDLQSPTKLSASGEDSTIP-QASRLNINVRSMGTGSPRSSSESHSPLYMHP 472
QY 267 -----KCG--APASTPYLLTSQ-----AQDTSGSPSERARPLPGSEFSCQNCVAGCSCG 316
Db 473 PKCTSGSOSPOHAEMCLHTAGPTFAEMGETQSEYSDSSCENGAFPCNECDRFESEAS 532
QY 317 LD-SLVPGDEDPYKQCRSSFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 375
Db 533 LKRLTLQTHSDPKYKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 592
QY 376 SRIHSGEKPKYKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 435
Db 593 TRIHSGEKPKYKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 614
QY 436 AHLRAHVLHTGKPYKCPCTCTRRHLQTLKSHVRIHTGKPYHCDPGLHFRHKSOLR 495
Db 615 AHLRAHVLHTGKPYKCPCTCTRRHLQTLKSHVRIHTGKPYHCDPGLHFRHKSOLR 674
QY 496 LHLROKHGAATNTKVHYHI 514
Db 675 LHLROKHGAATNTKVQYRV 693

RESULT 5
US-09-761-117-2
; Sequence 2, Application US/09761117
; Patent No. 6783945
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; Chaganti, Raju S.K.
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
; bcl-6
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/761,117
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US-Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 706 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-761-117-2

Query Match 39.3%; Score 1095.5; DB 4; Length 706;
Best Local Similarity 35.7%; Pred. No. 6.3e-76;

Matches 264; Conservative 60; Mismatches 144; Indels 271; Gaps 18;

QY 1 MGSPAPAGALGYVREFTTHSSDVLGNELRLGILTDVTLVGGQPLRAHKAVALIACS 60

Db 1 MASPA--DSCI---QFTHARDVLLNLRSLRDLTDVVIVVRSREQFRAHKTVMMAWR 54

QY 61 GFYSIFRGRAGVGVVLSLPGGPEARGFAPLDDPMYTSRLRLSPATAVLAATAATYQLM 120

Db 55 GLFYSIFTDQKCNLSVINLDPINPEGFCILDDPMYTSRLNLRGNIMAVNATANYQLM 114

QY 121 EHVQACHRFIOAS----- 134

Db 115 EHVVDTCRKFKIKASEAEAMVSAIKPPREBFNLSRMLMPQDIMAYRGREVVENNLPLRSAPG 174

QY 135 -----YEPLGIS----- 141

Db 175 CESRAFPALSYGLSTPSPASYSNMYSHLPVSSLLFSDEFRDVRMPVANPFPKRALPCDS 234

QY 142 -----LRLE----- 145

Db 235 ARPVPGEYSRPTLEVPNSVCHSNISYKETIPEARSMDHYSVAEGLKPAAPSARNAPYF 294

QY 146 -----EAEPT-----PPTAP-----PPGSRSEGHDPPTES-----RSC 177

Db 295 PCOKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSPKSDCQNSPTEACSSKNAC 354

QY 178 -----SQGPPSPASPPKACNWKYKIVLNS-----QASOAGSLVGRSSGQPCP 223

Db 355 ILQSGSPPAKSPPTDPKACSWKYKIVLNSLNQNAKPGPQAEGLRLSPRAYTAPPAC 414

QY 224 QARLPBGDEASSSSSSSSEEGIPQPSRL-----SPTAATVQF----- 266

Db 415 QPPM-EPENLDLQSPTKLSASGEDSTIP-QASRLNINVRSMGTGSPRSSSESHSPLYMHP 472

QY 267 -----KCG--APASTPYLLTSQ-----AQDTSGSPSERARPLPGSEFSCQNCVAGCSCG 316

Db 473 PKCTSGSOSPOHAEMCLHTAGPTFAEMGETQSEYSDSSCENGAFPCNECDRFESEAS 532

QY 317 LD-SLVPGDEDPYKQCRSSFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 375

Db 533 LKRLTLQTHSDPKYKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 592

QY 376 SRIHSGEKPKYKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 435

Db 593 TRIHSGEKPKYKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRPANLKH 614

QY 436 AHLRAHVLHTGKPYKCPCTCTRRHLQTLKSHVRIHTGKPYHCDPGLHFRHKSOLR 495

Db 615 AHLRAHVLHTGKPYKCPCTCTRRHLQTLKSHVRIHTGKPYHCDPGLHFRHKSOLR 674

QY 496 LHLROKHGAATNTKVHYHI 514

Db 675 LHLROKHGAATNTKVQYRV 693

RESULT 6
PCT-US94-08669-2
; Sequence 2, Application PC/TUS9406669
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Chaganti, R.S.K.
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS

; TITLE OF INVENTION: bcl-6
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10112
 ; COMPUTER READABLE FORM: disk
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/06669
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA: US 08/074,967
 ; APPLICATION NUMBER: 09-JUN-1993
 ; FILING DATE: 09-JUN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 43771-A-PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 977-9550
 ; TELEFAX: (212) 664-0525
 ; TELEX: 422523 COOP UI
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 706 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US94-06669-2

Query Match 39.38; Score 1095.5; DB 5; Length 706;
 Best Local Similarity 35.78; Pred. No. 6.3e-76; Indels 271; Gaps 18;
 Matches 264; Conservative 60; Mismatches 144;

QY 1 MGSPAPEGALGVYREFTSSDVLGNLNLRLGILTDVTLVGGQPLRAHKAVALIACS 60
 DB 1 MASPA--DSCI-----QTRHARDVLLNLRSLDILTDVVIVSRQFRAHKTVMW 54
 QY 61 GFFYSIFRGRAGVDVLSLPGGPEARGFAPLLDFMTSLRLSPATAPAVLAATYLOM 120
 DB 55 GLFYSIFTDLKCNLSVINLDPINPEGFCILLDFMTSLRLNREGNIMAVMATYLOM 114
 QY 121 EHVVOACHRFQAS----- 134
 DB 115 EHVVDTCRKFKTAKAEAEVMSAIKPPREFLNSRLMPQDINAYRGREVVNNLPLRSAPG 174
 QY 135 -----YEPLGIS----- 141
 DB 175 CESRAFAPSLYGLSTPPASYSMTSHLPVSSLLFSDEFRDVRMPVANPPKERALPCDS 234
 QY 142 -----LRPL----- 145
 DB 235 ARPVPGEYSRTELVSPNVCHSNITYSPKETIPEEARSMDHYSVAEGLKPAAPARNAPYF 294
 QY 146 -----EAPPT-----PPTAP-----PPGSPRSEGHDPPTES-----RSC 177
 DB 295 PCDKASKEEERFSSDEIALHFEPNAPLNRLKGLVSPQSPKQDCQPNPTEACSSKNAC 354
 QY 178 ----SQGPPSPADPKACNKKYKIVLNS-----QASQAGSLVGRSSGQPCP 223
 DB 355 ILOGSGSPAPKSPPTDFKACSKWKYKIVLNSLNQNAKGGPEQAEGLSLSPRAYTAPAC 414
 QY 224 QARLPDGEASSSSSSSSSSSEEGPIFGQSRL-----SPTAATVQF----- 266
 DB 415 QPPM-EPENLDLQSPTKLSASGEDSTIP-QASRLNINVRSMGTGSPRSSSESHSPLMYHP 472

QY 267 ----KCG--APASTPYLTQS-----AQDTSGSPSERARPLPGSEFFSQNCEAVAGCSSG 316
 DB 473 PKCTSCGSPQHAEMCLHTAGPTFAEEMGETQSEYSSDSCENGAFECNDCRFSERAS 532
 QY 317 LD-SLVPGDEDPKPKQLCRSSFRYKGNLASHRTVHTGKPYHCSICGARFNRPNLAKTH 375
 DB 533 LKHHTLQTHSDPKPKCDRCQASFRYKGNLASHRTVHTGKPYRNCICGAQFNPNLAKTH 592
 QY 376 SRIHSGEKPKCETCGSRFVQVRSQPPSGFQGPARGGVGQKGGPCSSQRODLKSPSOV 435
 DB 593 TRIHSGEKPKCETCGARFV-----QV 614
 QY 436 AHLRAHVLHTGKPYPCPTCTGTRFRHLQTLKSHVRIHTGKPYHCDPGLHPRHKSQRL 495
 DB 615 AHLRAHVLHTGKPYPCPTCTGTRFRHLQTLKSHVRIHTGKPYHCEKCNLHPRHKSQRL 674
 QY 496 LHLROKHGAANTNTKVHYHI 514
 DB 675 LHLROKHGAINTNTKVQYRV 693

RESULT 7
 US-09-063-035-2
 ; Sequence 2, Application US/09063035
 ; Patent No. 6160091
 ; GENERAL INFORMATION:
 ; APPLICANT: PEUKERT, Karen; HAENEL, Frank; and EILERS,
 ; APPLICANT: Martin
 ; TITLE OF INVENTION: Myc-binding zinc finger proteins,
 ; TITLE OF INVENTION: their preparation and their use
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Keil & Weinlauf
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage
 ; COMPUTER: IBM AT-compatible, 80486 processor
 ; OPERATING SYSTEM: MS-DOS version 6.1
 ; SOFTWARE: Wordperfect version 8.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/063,035
 ; FILING DATE: 21-APR-1998
 ; CLASSIFICATION: 514
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 803 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-09-063-035-2

Query Match 20.78; Score 576.5; DB 3; Length 803;
 Best Local Similarity 29.44; Pred. No. 5.5e-36;
 Matches 180; Conservative 64; Mismatches 205; Indels 163; Gaps 20;

QY 16 EFTFHSDDLGNLNLRLGILTDVTLVGGQPLRAHKAVALIACSGFFYSIFRGRAG-VG 74
 DB 2 DFFQHSQHVLEUQNLQQLGLLCCDCTFVDGVGHFAKHAVALACSEYFKMLFVDQKDVH 61
 QY 75 VDVLSPGGPEARGFAPLLDFMTSLRLSPATAPAVLAATYLOMEHVVOACHRFQAS 134
 DB 62 LDI-----SNAAGLGQMLEFMYTAKLSLSPENVDDVLAVATFLQMODIITACH----- 109
 QY 135 YEPLGISRLPLEAPPTPP-----TAPPGSPRSEGHDPPTESRSCSGPSPASPD 188
 DB 110 -----ALKSL-ARPATSPGNAEALATEGGDKAKBEKVATSTLSRLEQAGRSTPGP- 161
 QY 189 PKACNKKYKIVLNSQASQAGSLVGRSSGQPCQARLP-----SGDEASSSSSSSSSSSS 245
 DB 162 --SRLKEERGQQAASGAEBQTEKADAPREPPFVELKPDPTSGMAAEAEALSESSE 219

Db 178 PGT--SDLGFLDMTNGAALANSNGIAGSMQPEEA-----ARAGAAIA 221
Qy 215 ERSSGQPCQA-RLPSSGDEASS--SSSSSSSEEGPIPCQSRSLSTPTAATVQKCA 270
Db 222 QASLPLVPGVDRPLMWAGPLSPOLLTSPPFSVASSAPPLTGKRGGRPRKAN-----274
Qy 271 PASTPYLLTSAQDTSQSPS--ERARPLP-----GSEFFSCQNC-----EAVAGCSS--GLD 318
Db 275 -----LLDSM-----FGSPGGLREAGILPCGLCGKVFDTANRLRQHEAQHGVTSLQLGYI 324
Qy 319 SLVP-----GSDKPYK-----COLCRSSFRYKGNLASHRTVHTGKPYH 358
Db 325 DLPPRLGGLPTISEDPGPKRSRTRKQVACEICGKIFRDVYHLNRHKLHSHSGEKPY 384
Qy 359 CSICGARENPNANKTHSRHSIGE--KPYKCTCGSRFVQVRSQPPSGFGKPARGGVQ 416
Db 385 CPVCLRFKPKRDMYSYHVRSHDGSVGKPYICQSCGKF-----422
Qy 417 KGGFCSSQDQLKSPSOVAHLRAHV-LIHTGKPYPCPTCTGTRPHLQTLKSHVRIH-- 473
Db 423 -----SRPDHLNGHIKQVHTSERPHKQTCNASFATDRLRSLHSLACHED 466
Qy 474 -----473
Db 467 KVPQCVGKYLRAAYMADHLKXHSSEPSNFCISCNREGQKCSHODPIESSDSYGLSDAS 526
Qy 474 -----TGEKPYHCDPCGLHFRHKSQRLHLRQKHGAAT 506
Db 527 DLKTPKQANGSFCDMAVPKKNMESDGKKYPCPCGCGSFRSKSYLNKHIQ-----579
Qy 507 NTKVHYHILGGP 518
Db 580 --KVHVRALGGP 589

RESULT 10

US-09-949-016-10889
; Sequence 10889, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10889
; LENGTH: 641
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10889

Query Match 16.4%; Score 457.5; DB 4; Length 641;
Best Local Similarity 25.0%; Pred. No. 5.7e-27;
Matches 168; Conservative 56; Mismatches 195; Indels 253; Gaps 25;

Qy 7 PEGALGYVREFTRHSSDVLGNLNLRLR-GILTDVLLVGGQPLRAHKAVALIACSGFFYS 65
Db 11 PSGC--YTYQVSRHSTEMLNLRNQRKNGRFGCDVLLRVGDESFPFAHRAVLACSEYFES 68
Qy 66 IFRGRAGVG-----VDV-----LSLPGP-----EARGFAPLDFMTYSLRL 103
Db 69 VFSALQDGGADGGPADVGGATAPGGAGGSGRELMHTSSKVFVGDILDFAYTSRIIV 128

Qy 104 SPATAPAVLAATYLOMEHVVOACHRFQIAQSYELGILSRLEAPPTPTAP-----P 157
Db 129 RLESPPELMTAAKFLLMRSVIEICQEVIKQS-----NVQILVPPARADIMLPRP 177
Qy 158 PGSPRSEGHDPPTESRSC--SQGPPSPASPPDKACNWKYKIVILNQSQAQSLVG 214
Db 178 PGT--SDLGFLDMTNGAALANSNGIAGSMQPEEA-----ARAGAAIA 221
Qy 215 ERSSGQPCQA-RLPSSGDEASS--SSSSSSSEEGPIPCQSRSLSTPTAATVQKCA 270
Db 222 QASLPLVPGVDRPLMWAGPLSPOLLTSPPFSVASSAPPLTGKRGGRPRKAN-----274
Qy 271 PASTPYLLTSAQDTSQSPS--ERARPLP-----GSEFFSCQNC-----EAVAGCSS--GLD 318
Db 275 -----LLDSM-----FGSPGGLREAGILPCGLCGKVFDTANRLRQHEAQHGVTSLQLGYI 324
Qy 319 SLVP-----GSDKPYK-----COLCRSSFRYKGNLASHRTVHTGKPYH 358
Db 325 DLPPRLGGLPTISEDPGPKRSRTRKQVACEICGKIFRDVYHLNRHKLHSHSGEKPY 384
Qy 359 CSICGARENPNANKTHSRHSIGE--KPYKCTCGSRFVQVRSQPPSGFGKPARGGVQ 416
Db 385 CPVCLRFKPKRDMYSYHVRSHDGSVGKPYICQSCGKF-----422
Qy 417 KGGFCSSQDQLKSPSOVAHLRAHV-LIHTGKPYPCPTCTGTRPHLQTLKSHVRIH-- 473
Db 423 -----SRPDHLNGHIKQVHTSERPHKQTCNASFATDRLRSLHSLACHED 466
Qy 474 -----473
Db 467 KVPQCVGKYLRAAYMADHLKXHSSEPSNFCISCNREGQKCSHODPIESSDSYGLSDAS 526
Qy 474 -----TGEKPYHCDPCGLHFRHKSQRLHLRQKHGAAT 506
Db 527 DLKTPKQANGSFCDMAVPKKNMESDGKKYPCPCGCGSFRSKSYLNKHIQ-----579
Qy 507 NTKVHYHILGGP 518
Db 580 --KVHVRALGGP 589

RESULT 11

US-09-949-016-10890
; Sequence 10890, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10890
; LENGTH: 641
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10890

Query Match 16.4%; Score 457.5; DB 4; Length 641;
Best Local Similarity 25.0%; Pred. No. 5.7e-27;
Matches 168; Conservative 56; Mismatches 195; Indels 253; Gaps 25;

Qy 7 PEGALGYVREFTRHSSDVLGNLNLRLR-GILTDVLLVGGQPLRAHKAVALIACSGFFYS 65
Db 11 PSGC--YTYQVSRHSTEMLNLRNQRKNGRFGCDVLLRVGDESFPFAHRAVLACSEYFES 68


```
; ORGANISM: Human
US-09-949-016-10273

Query Match      15.4%; Score 428.5; DB 4; Length 678;
Best Local Similarity 25.4%; Pred. No. 1e-24;
Matches 156; Conservative 63; Mismatches 235; Indels 159; Gaps 19;

QY 20 HSSDVLGNLRLRGLTDTVLLVGGQPLRAHKAVALIACGFFYSIF-RGRAGVGVNVL 78
Db 21 HPTGLLCKANQRLAGTLCDDVIMVDSQEFHARIVLACTSKMFEILPHRNSQHYTLDFL 80

QY 79 SLPGGPEARGAPLDFMYTGRRLSPATAPAVLAATYLOMEHVQACHRF---IQASY 135
Db 81 S-----PKTFQILLEYATATLOAKAEDLDLLVAAILREYLEEQCLMLETIOASD 134

QY 136 E----- 136
Db 135 DNDTEATWADGAAEEEDRKARYLNIFISKHSSEGSYASVAGOSLPGMVDQSPSVST 194

QY 137 -----PLGISLRPLEAEPPTPTAPPSPRRSEGHPPPTTESRSCS 178
Db 195 SFGLSAMPTKAVDSLMTIGOSLLQGLTLOPPAGPEETLAGGRHGPVAVKTEMQVD 254

QY 179 QGPPSPASDPKACNMWKYKIVLNSQASQAGSLVGRSSGQPCQAR-----LPS 229
Db 255 E---VPSQDSPCAAE-----SSISGMDKVEERKGEPTPTTRSSVITSARELHY 302

QY 230 GDEASS-----SSSSSSSSSEEGPIPGPQGRSLSPATAVQFKGAPASTPYLLTS--- 280
Db 303 GREESAQVPPPAEAGQAPTGRPEHPAPPPEKHLGIYSLVPHKADAVLSMPSSVTSGLH 362

QY 281 --QAQDTSGSPSERARPLGS-----EFFSCNCEAVAGSSGLDVLVPDDEKPKYKCOLC 334
Db 363 VQPALAVSNDFTYCGLLPQGGFIQELFS-KLGLAVGMS--ESRTIGEO-----CVC 414

QY 335 RSSFRYKGNLASHRTVHTGKEPYHSGICGARNRNPANLKTSHRIS-GEKPYKBTGCSR 393
Db 415 GVLEPDNEAVEQHRKLSGKMTYGCGLCKRFLDSLRLMHLAHSAGAKAFVCDQCGAQ 474

QY 394 F-----VQVRSQPPSG-----FQKPA-----RGVQKQKGFSSQQRDLKS 430
Db 475 FSKEDALETHRTHTGTDMVAFCLLCGRKRFQASALQQHMEHVAGVRSYICSECNRTF-- 532

QY 431 PPSQVAHLRAHVLITGKEPYCPCTGTFRHLQTLKSHVRIHTGKPYHCDPCGLHFRH 490
Db 533 -PSHTA-LKRHLRSHTDHPYCEFCGFCRDESLKSHKRIHTGKPYECNGCGKFKSL 590

QY 491 KSQRLHLRQKHG 503
Db 591 KQLETHRVHTG 603

RESULT 14
US-09-949-016-8149
; Sequence 8149, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8149
; LENGTH: 769
```

```
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8149

Query Match      15.1%; Score 422.5; DB 4; Length 769;
Best Local Similarity 28.8%; Pred. No. 3.5e-24;
Matches 109; Conservative 46; Mismatches 140; Indels 83; Gaps 9;

QY 205 QASQAGSLVGRSSGQPCQARLPSPGDEASSSSSSSSSSSSSEEGPIPGPQGRSLSPATAV 264
Db 295 QGNECEAFNDSSLELHKQVHLCKSPACSTHEKDTSYSS--GIPVQSVRTGKKRYW 351

QY 265 QFKCGAPAS--TPYLLTSQAQDTSGSPSERARPLPGSFFSCNCEAVAGSSGLDVLSP- 322
Db 352 CHECGKFSQSSNLQTHRVHTGKPE-----YTCECGKSFNQSHLYAHLPI 399

QY 323 --GDE-----DKPYKCOLCRSSFRYKGNLASHRTVHTGE 354
Db 400 HTGKPYRCDSGCGFGRSRDNLNHCVRHTGKEPYKCEVCGKGTQSRHLQAHRIHTGE 459

QY 355 KPYHCSTICGARFNRPANLKTSHRIHSGEKKPYKCTCGSRF-----VQVRSQPPS 403
Db 460 KPYKCGCGKGFSSCSNLHTQHVHTGKPYKCDCECGKCFSLFNLHSHQHVHTGKPE-- 517

QY 404 GFQKPARGGVQKGGFCSSQORDLKSP-----SQVAHLRAHVLITGKPYPC 453
Db 518 -YKCECGKGFSSASSFSQSHQHVHTGKPPFCNVCCKGFSQSSYFQAHQHVHTGKPYK 576

QY 454 PTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQRLH-----LR 499
Db 577 EVCKGRFNWLSLNLNHCVRHTGKPYKCECGKGFSSQASNLQAHQSVHTGKPFKCDACQ 636

QY 500 QKHGAATNTKVHYHILGG 517
Db 637 KRFSQASHLQAHQHVHTG 654

RESULT 15
US-09-538-092-1356
; Sequence 1356, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 1356
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q16600
US-09-538-092-1356

Query Match      15.1%; Score 421.5; DB 4; Length 458;
Best Local Similarity 30.5%; Pred. No. 2.2e-24;
Matches 107; Conservative 32; Mismatches 129; Indels 83; Gaps 10;

QY 176 SCGQPPSPASDPKACN-----WKYKIVLNSQA-----SQAGSLVGRSS 218
Db 137 TCQNG-QLKESLDLPIDCNCKDIHWK--SQVSCSQQRGHTEKPCDHNNCKILNTSPD 193

QY 219 QGPCQARLPSPGDEASSSSSSSSSSSEEGPIPGPQGRSLSPATAVQFKGAPASTPYLL 278
```

Db	194	CHPYEKIHTAEKQYEGSQCKNFSQSSE-----LL	223
Qy	279	TSQAQDTSGSPSERARPLPGSEFFSQCNC-EAVAGCSSGLDSLVPDDEKPKYKCOLCRSS	337
Db	224	LHQRDHTEKP-----YKCEQCGKGFTRSSLLIHQAVHTDEKPKYCEKCGK	271
Qy	338	FRYKGNLASHRTVHTGEKPYHCSI CGARFNRPANLKTSHRSIHSGEKPYKCTCGSRFVQ-	396
Db	272	FTRSSLLLIHHAHVHTGEKPYKCDKCGKGFSSQSKLHIHQHVHTGEKPYECECGMSFSQR	331
Qy	397	-----VRSOPPSGFQKPARGGVQKGGFCSSORQDLKSPPSQVAHLRAHVLHTGEKPY	452
Db	332	SNLHIHQDHTG--ERPYKCGECGKGF-----SOSNLHIHRCIHTGEKPYQ	376
Qy	453	CPTCGTRFRHLQTLAKSHVRIHTGEKPYHCDPCGLHFRHKSQLRLHLRQKHG	503
Db	377	CYECGKGFSSQSDLRHLRVHTGEKPYHCGKCGKGFSSQSKLIIHQRVHTG	427

Search completed: August 9, 2005, 14:06:58
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 13:59:52 ; Search time 160 Seconds
(without alignments)

Title: US-09-815-370-4
 perfect score: 2789
 Sequence: 1 MGSPAAPEGALGVYREFTRH.....ROKHGAATNTKVYHILGSP 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1752860 seqs, 390397842 residues

Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database : Published Applications AA.*
1: /cgm2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgm2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgm2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgm2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgm2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgm2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgm2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgm2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgm2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgm2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgm2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgm2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgm2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgm2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgm2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgm2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgm2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp.*
18: /cgm2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
19: /cgm2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pcp.*
20: /cgm2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
21: /cgm2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
22: /cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match	%				
1	2789	100.0		518	10	US-09-815-379-4	Sequence 4, Appli
2	2541.5	91.1		479	15	US-10-295-027-16	Sequence 16, Appli
3	2541.5	91.1		479	15	US-10-211-462-24	Sequence 24, Appli
4	1128.5	40.5		706	16	US-10-755-889-66	Sequence 66, Appli
5	1126.5	40.4		706	16	US-10-408-765A-1988	Sequence 1988, Ap
6	1122.5	40.2		706	16	US-10-755-889-18	Sequence 18, Appli
7	1095.5	39.3		706	9	US-09-107-058-2	Sequence 2, Appli
8	1095.5	39.3		706	9	US-09-761-117-2	Sequence 9, Appli
9	575.5	20.6		803	15	US-10-441-854-9	Sequence 12, Appli
10	543.5	19.5		765	18	US-10-486-977-12	Sequence 4, Appli
11	509.5	18.3		539	17	US-10-965-998-4	

12	483.5	17.3	697	16	US-10-408-765A-2256	Sequence 2256, Ap
13	482.5	17.3	697	16	US-10-723-8620-4408	Sequence 4408, Ap
14	482	17.3	610	15	US-10-221-6325-10	Sequence 10, Appl
15	469.5	16.8	711	14	US-10-034-934-90	Sequence 90, Appl
16	457.5	16.4	641	14	US-10-164-339-17	Sequence 17, Appl
17	447	16.0	687	16	US-10-408-765A-1905	Sequence 1905, Ap
18	434	15.6	614	16	US-10-408-765A-1905	Sequence 13, Appl
19	431.5	15.5	535	15	US-10-264-049-3016	Sequence 3016, Ap
20	428.5	15.4	616	16	US-10-478-019-169	Sequence 169, App
21	428.5	15.4	678	15	US-10-296-615-1316	Sequence 1316, Ap
22	428	15.3	516	9	US-09-974-298-118	Sequence 118, App
23	428	15.3	516	9	US-10-094-749-1976	Sequence 1976, Ap
24	426.5	15.3	127	10	US-09-764-891-4337	Sequence 4337, Ap
25	423.5	15.2	548	15	US-10-085-196-436	Sequence 136, App
26	423.5	15.2	670	15	US-10-108-260A-3103	Sequence 3103, Ap
27	423	15.2	498	15	US-10-094-749-1657	Sequence 1657, Ap
28	422	15.1	464	15	US-10-108-260A-3590	Sequence 3590, Ap
29	421.5	15.1	458	16	US-10-719-993-450	Sequence 450, App
30	421.5	15.1	622	15	US-10-092-900A-332	Sequence 332, App
31	420	15.1	1077	16	US-10-408-765A-2871	Sequence 2871, Ap
32	419.5	15.0	458	16	US-10-719-993-448	Sequence 448, App
33	419.5	15.0	458	16	US-10-719-993-449	Sequence 449, App
34	419.5	15.0	488	9	US-09-840-787-17	Sequence 17, Appl
35	419.5	15.0	819	15	US-10-108-260A-3302	Sequence 3302, Ap
36	418.5	15.0	488	16	US-10-408-765A-2340	Sequence 2340, Ap
37	418.5	15.0	658	16	US-10-475-681-10	Sequence 10, Appl
38	418.5	15.0	659	16	US-10-475-681-6	Sequence 6, Appl
39	417.5	15.0	595	16	US-10-203-052B-8	Sequence 8, Appl
40	417	15.0	406	9	US-09-764-864-976	Sequence 976, App
41	417	15.0	638	16	US-10-723-860-1613	Sequence 1613, Ap
42	416.5	14.9	519	15	US-10-094-749-1972	Sequence 1972, Ap
43	415.5	14.9	803	16	US-10-408-765A-1049	Sequence 1049, Ap
44	415	14.9	556	15	US-10-104-047-3914	Sequence 3914, Ap
45	413.5	14.8	799	15	US-10-104-047-3929	Sequence 2929, Ap

ALIGNMENTS

RESULT 1
 US-09-815-379-4
 ; Sequence 4, Application US/09815379
 ; Publication No. US20030073613A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RASTELLI, LUCA
 ; APPLICANT: GERRITSEN, MARY
 ; TITLE OF INVENTION: ANGIOGENESIS ASSOCIATED PROTEINS AND NUCLEIC ACIDS
 ; TITLE OF INVENTION: ENCODING THE SAME
 ; FILE REFERENCE: 10716/35
 ; CURRENT APPLICATION NUMBER: US/09/815,379
 ; CURRENT FILING DATE: 2001-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,134
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 518
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-379-4

	Query Match	100.0%	Score 2789;	DB 10;	Length 518;
	Best Local Similarity	100.0%;	Pred. No. 2.2e-162;		
	Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
y	1	MGSPAAPGALGYVREFTRHSSDVLGNLNEURLRGILTDVTLVGGQPLRAHKAVLIACS	60		
b	1	MGSPAAPGALGYVREFTRHSSDVLGNLNEURLRGILTDVTLVGGQPLRAHKAVLIACS	60		
y	61	GFYYSIFRGRAGVGDVLSLPGGPARGFAPLLDFMITSRLRLSPATAPAVLAATYLOM	120		
b	61	CFEYYSIFRGRAGVGDVLSLPGGPARGFAPLLDFMITSRLRLSPATAPAVLAATYLOM	120		


```
QY 121 EHVQACHRFQASVEPLGSLRPLEAEPPTPTAPPFGSPRRSEGHDPPTESRSCSQ 180
DB 121 EHVQACHRFQASVEPLGSLRPLEAEPPTPTAPPFGSPRRSEGHDPPTESRSCSQ 180
QY 181 PPSASDPKACNWKYKIVLNSQASGSLVGRSSGQPCQARLPSPGDEASSSSSS 240
DB 181 PPSASDPKACNWKYKIVLNSQASGSLVGRSSGQPCQARLPSPGDEASSSSSS 240
QY 241 SSSSEEGPIPGQSRLSPTAATVQFKGAPASTPYLLTSAQDTSGPSERARPLPGE 300
DB 241 SSSSEEGPIPGQSRLSPTAATVQFKGAPASTPYLLTSAQDTSGPSERARPLPGE 300
QY 301 FFSQNCBAVAGCSGLDLPVGDDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360
DB 301 FFSQNCBAVAGCSGLDLPVGDDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360
QY 361 ICGARFNRPANLKTSHRSIHSGEKPKYKCTCGSRFVQVRSOPPSGFGQKPGVQKGGF 420
DB 361 ICGARFNRPANLKTSHRSIHSGEKPKYKCTCGSRFVQVRSOPPSGFGQKPGVQKGGF 420
QY 421 CSSORODLKSPPSOVAHLRAHVLHTGKPYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
DB 421 CSSORODLKSPPSOVAHLRAHVLHTGKPYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
QY 481 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
DB 481 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
```

RESULT 2

```
US-10-295-027-16
; Sequence 16, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
```

```
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-16
Query Match 91.1%; Score 2541.5; DB 15; Length 479;
Best Local Similarity 92.5%; Pred. No. 2.7e-147;
Matches 479; Conservative 0; Mismatches 0; Indels 39; Gaps 2;
QY 1 MGSPAAPEGALGVYREFTRHSSDVLGNLNLRLGILTDVTLVGGOPLRAHKAVALIACS 60
DB 1 MGSPAAPEGALGVYREFTRHSSDVLGNLNLRLGILTDVTLVGGOPLRAHKAVALIACS 60
QY 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLDFMYTSLRLSPATAPAVLAATYIQM 120
DB 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLDFMYTSLRLSPATAPAVLAATYIQM 120
QY 121 EHVQACHRFQASVEPLGSLRPLEAEPPTPTAPPFGSPRRSEGHDPPTESRSCSQ 180
DB 121 EHVQACHRFQASVEPLGSLRPLEAEPPTPTAPPFGSPRRSEGHDPPTESRSCSQ 180
QY 181 PPSASDPKACNWKYKIVLNSQASGSLVGRSSGQPCQARLPSPGDEASSSSSS 240
DB 181 PPSASDPKACNWKYKIVLNSQASGSLVGRSSGQPCQARLPSPGDEASSSSSS 240
QY 241 SSSSEEGPIPGQSRLSPTAATVQFKGAPASTPYLLTSAQDTSGPSERARPLPGE 300
DB 241 SSSSEEGPIPGQSRLSPTAATVQFKGAPASTPYLLTSAQDTSGPSERARPLPGE 300
QY 301 FFSQNCBAVAGCSGLDLPVGDDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 359
DB 301 FFSQNCBAVAGCSGLDLPVGDDEKPKYKQLCRSSFRYKGNLASHRTVHTGKPYHCS 359
QY 361 ICGARFNRPANLKTSHRSIHSGEKPKYKCTCGSRFVQVRSOPPSGFGQKPGVQKGGF 420
DB 361 ICGARFNRPANLKTSHRSIHSGEKPKYKCTCGSRFVQVRSOPPSGFGQKPGVQKGGF 420
QY 421 CSSORODLKSPPSOVAHLRAHVLHTGKPYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
DB 421 CSSORODLKSPPSOVAHLRAHVLHTGKPYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
QY 481 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 518
DB 442 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHILGGP 479
```

RESULT 3

```
US-10-211-462-24
; Sequence 24, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 479
```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1988

Query Match      40.4%; Score 1126.5; DB 16; Length 706;
Best Local Similarity 35.0%; Pred. No. 1.2e-60;
Matches 270; Conservative 48; Mismatches 117; Indels 337; Gaps 16;

1 MGSPAAPGALGVYREFTRHSSDVLGNLNELRGLTDTVLLVGGQPLRAHKAVALIACS 60
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1 MASPA--DSCI---QFTRHASDVLLNLRSLRDLTDVVIVVSREQFRAHKTVMACS 54
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

61 GFYISIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATPAVLAATYLOM 120
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
55 GLFYISIFTDLKCNLSVINLDPINPEGFCILLDFMYTSRLNLRGNIMAVMATYLOM 114
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

121 EHVQOACHRFIOAS----- 134
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

115 EHVDTCKRFKASEAEWVAIKPPREFLNSRMLMPQDIMAYRGVVENNLPURSAPG 174
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

135 -----YPLGIS----- 141
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

175 CESRAFAPSLYGLSTPPASYSMYSHLPVSSLLFSDEEFDVRVMPVANPFPKERALPCDS 234
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

142 -----LRPL----- 145
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

235 ARPVPGEYSRPTLEVPSPNVCHSNISYSPKETIPEEARSMDHYSAEGLKPAAPSARNAPYF 294
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

146 -----EAEPT-----PPTAP-----PGSPRSEGHDPPTESRS----- 176
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

295 PCDKASKEERPSSEDEIALHFEPNAPLNRLKGLVSPQSKDCQPNSTPESCKNAC 354
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

177 ---CSGPPSPASPPKACNWKYKIVLNSQASQAGSLVGRSSGQPCQARLPSGDEA 233
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

355 ILQASGPPAKSPTDPKACNWKYKIVLNSQNAKPGGPEQAEGLRLSPRAYTAPAC 384
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

234 SSSSSSSSSSEEGIPGQSRLSPTAATVQKCAP-----ASTPYLLTSQAQDT- 285
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

385 ----SLNQNAKPEGEQAELGRLSPRAYTAPACQPMPEPENLDLQSPTKLSASGEDST 439
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

286 -----SGSP---SERAPL----- 296
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

440 IPOASRLNINVRSMTGSPPRSSESHSPLYMHPKCTSCGSPQHAEMCLHTAGTPTPE 499
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

297 ----PGSEF-----FSCNCAVAGCSSLGLD-SLVPGDDEKPKYKQLCRSSFYK 342
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

500 EMGETQSEYSDSCENGAFNCECDRCRSEELAKRHTLTQTHSDPKYKCDRCQASFRYK 559
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

343 NLASHRTVHTGKPYKHCSTCGARFNRPANLKTHTSRHSGEKPKYKCTGSRFVQVRSPP 402
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

560 NLASHKTVHTGKPYKRCNICGAQFNRPANLKTHTSRHSGEKPKYKCTGSRFVQVRSPP 612
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

403 SFGQKPARGGVGQGGFCSSQDLSKPPSOVAHLRAHVLHTGKPYKCTGCTTRFRH 462
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

613 -----QVAHLRAHVLHTGKPYKCTGCTTRFRH 641
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

463 LQTLKSHVRIHTGKPYKCHDCPLHFRHKSQRLRLRQKHGAATNTKVHYHI 514
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

642 LQTLKSHLRIHTGKPYKCHCEKRLHFRHKSQRLRLRQKHGAITNTKVQYRV 693
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 6
US-10-755-889-18
; Sequence 18, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
```

```
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-18

Query Match      40.2%; Score 1122.5; DB 16; Length 706;
Best Local Similarity 36.3%; Pred. No. 2.1e-60;
Matches 268; Conservative 58; Mismatches 142; Indels 271; Gaps 17;

1 MGSPAAPGALGVYREFTRHSSDVLGNLNELRGLTDTVLLVGGQPLRAHKAVALIACS 60
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1 MASPA--DSCI---QFTRHASDVLLNLRSLRDLTDVVIVVSREQFRAHKTVMACS 54
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

61 GFYISIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATPAVLAATYLOM 120
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
55 GLFYISIFTDLKCNLSVINLDPINPEGFCILLDFMYTSRLNLRGNIMAVMATYLOM 114
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

121 EHVQOACHRFIOAS----- 134
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

115 EHVDTCKRFKASEAEWVAIKPPREFLNSRMLMPQDIMAYRGVVENNLPURSAPG 174
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

135 -----YPLGIS----- 141
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

175 CESRAFAPSLYGLSTPPASYSMYSHLPVSSLLFSDEEFDVRVMPVANPFPKERALPCDS 234
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

142 -----LRPL----- 145
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

235 ARPVPGEYSRPTLEVPSPNVCHSNISYSPKETIPEEARSMDHYSAEGLKPAAPSARNAPYF 294
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

146 -----EAEPT-----PPTAP-----PGSPRSEGHDPPTESRS----- 176
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

295 PCDKASKEERPSSEDEIALHFEPNAPLNRLKGLVSPQSKDCQPNSTPESCKNAC 354
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

177 ---CSGPPSPASPPKACNWKYKIVLNSQASQAGSLVGRSSGQPCP 223
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
355 ILQASGPPAKSPTDPKACNWKYKIVLNSQNAKPGGPEQAEGLRLSPRAYTAPAC 414
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

224 QARLPDGEASSSSSSSSEEGIPGQSRLSPTAATVQF-----SPTAATVQF----- 266
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

415 QPPM-EPENLDLQSPTKLSASGEDSTIP-QASRLNINVRSMTGSPPRSSESHSPLYMHP 472
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

267 ----KCG--APASTPYLLTSQ----AODTSGSPSERARPLPGSEFFSCQNCQAVAGCSSG 316
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

473 PKTSCGSPQHAEMCLHTAGTPTFAEMGETQSEYSDSCENGAFNCECDRCRSEAS 532
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

317 LD-SLVPGDDEKPKYKQLCRSSFYKGNLASHRTVHTGKPYKHCSTCGARFNRPANLKTHT 375
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
533 LKRLHTLQTHSDPKYKCDRCQASFRYKGNLASHKTHTGKPYKRCNICGAQFNRPANLKTHT 592
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

376 SRIHSGEKPKYKCTGSRFVQVRSQPPSGFGQCKPARGVGQGGFCSSQDLSKPPSOV 435
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
593 TRIHSGEKPKYKCTGSRFVQVRSQPPSGFGQCKPARGVGQGGFCSSQDLSKPPSOV 614
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

436 AHLRAHVLHTGKPYKCTGCTTRFRHLQTLKSHVRIHTGKPYKCHDCPLHFRHKSQRL 495
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
615 AHLRAHVLHTGKPYKCTGCTTRFRHLQTLKSHLRIHTGKPYKCHCEKRLHFRHKSQRL 674
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

496 LHLRQKHGAATNTKVHYHI 514
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
675 LHLRQKHGAITNTKVQYRV 693
  ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 7
US-09-107-058-2
; Sequence 2, Application US/09107058
; Patent No. US20010010922A1
; GENERAL INFORMATION:
```

APPLICANT: Dalla-Favera, Riccardo
 APPLICANT: Niu, Hui-Peng
 TITLE OF INVENTION: CLONING AND USES OF THE GENETIC
 TITLE OF INVENTION: LOCUS bcl-6
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,058
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 TELEX: 422523 COOP UI
 FILING DATE: 16-Jan-2001
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 706 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-107-058-2

Query Match 39.3%; Score 1095.5; DB 9; Length 706;
 Best Local Similarity 35.7%; Pred. No. 9.4e-59;
 Matches 264; Conservative 60; Mismatches 144; Indels 271; Gaps 18;
 1 MGSPAPEGALGVYREFTHSSDVLGNLNLRLGILTDVLLVGGQPLRAHKAVALIACS 60
 1 MASPA-DSCI-----QFTHARDVLLNLRSLRDLTDVIVVREQFRAHKTVMWR 54
 61 GFYSGIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLSLSPATAPVLAATYLOM 120
 55 GLFYSIPTDQKCNLSVINLDPINPEGFCILLDFMYTSRLNLRGNIMAVMATAYLOM 114
 121 EHVQACHRFIOAS----- 134
 115 EHVDTCRKFIKASEAEMVSAIKPREEFLNRMMLPQDIMAYRGVVENNPLRSAPG 174
 135 -----YELGIS----- 141
 175 CESRAPSLYGLSTPPASYSNYSHPVLSLLFSDSEDFRDMVPVANDPFRALPCDS 234
 142 -----LRPL----- 145
 235 ARPVPGEYSRPTLEVPNVCHSNISYSPKETIPEARSMDHYSVAEGLKPAAPSARNAPYF 294
 146 -----EAPPT-----PPTAP-----PPGSPRRSEGHDPPTES-----RSC 177
 295 PCDKASKEERPESEDEIALHFPNAPLNRLKGLVSPQSPKSDCPNPSPTACSSKNAC 354
 178 -----SGPSPASPDPKACNWKYKIVLNS-----QASQAGSLVGERSSQPCP 223
 355 ILQSGSPPAKSTDPDKASCWKYKIVLNSLNQNAKPGQPEQAEGLRSLSPRAYTAPPAC 414
 224 QARLPGDEASSSSSSSSSSSREGPIPGQSR-----SPTAATVQF----- 266
 415 QPPM-EPENLDQSPTKLSASGEDSTIP-QASRLNINVRMSVTGSPRSSSESHSLYMPH 472

267 ----KCG--APASTPYLLTSQ-----AQDTSGSPSERARPLPGSEFFSCNCEAVAGCSG 316
 473 PKCTSCGSPQHAEMCLHTAGPTFAEMGETQSEYSDSSCENGAFECNDCRFSSEAS 532
 317 LD-SLVPGDDEKPKCOLCRSSFRYKGNLASHRTVHTGKPYHCSICGARFNRRPANKLTH 375
 533 LKXHTLQTHSDKPKCDRCQASFRYKGNLASHRTVHTGKPYHCSICGARFNRRPANKLTH 592
 376 SRIHSGEKPKYKCTCGSRFVQVRSQPPSPGQKPARGGVQGGFCSSQRODLKSPPSQV 435
 593 TRIHSGEKPKYKCTCGARFV-----QV 614
 436 AHLRAHVLIHTGKPYKCTCGTRPRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQLR 495
 615 AHLRAHVLIHTGKPYKCTCGTRPRHLQTLKSHLRIHTGKPYHCEKCNLHFRHKSQLR 674
 496 LHLROKHGAATNTKVHYHI 514
 675 LHLROKHGAATNTKVQYRV 693

RESULT 8
 US-09-761-117-2
 ; Sequence 2, Application US/09761117
 ; Patent No. US20010012887A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dalla-Favera, Riccardo
 ; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/761,117
 ; FILING DATE: 16-Jan-2001
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US-Y
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525
 ; TELEX: 422523 COOP UI
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 706 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-761-117-2

Query Match 39.3%; Score 1095.5; DB 9; Length 706;
 Best Local Similarity 35.7%; Pred. No. 9.4e-59;
 Matches 264; Conservative 60; Mismatches 144; Indels 271; Gaps 18;
 1 MGSPAPEGALGVYREFTHSSDVLGNLNLRLGILTDVLLVGGQPLRAHKAVALIACS 60
 1 MASPA-DSCI-----QFTHARDVLLNLRSLRDLTDVIVVREQFRAHKTVMWR 54
 61 GFYSGIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLSLSPATAPVLAATYLOM 120

Db 55 GLFYISFTDQKCNLSVINLDPETINPEGFCILLDFMYTSRLNREGNIMATAMYLQ 114
Qy 121 EHVQACHRFIOAS
Db 115 EHVDTCKRFKASEAEMVSAIKPPREFFLNSMLMPODIMAYRGREVENNLPLRSAPG 174
Qy 135 -----YPLGIS----- 141
Db 175 CESRAFAPSLYGLSTPPASYSMYSHLPVSSILLFSDERFDRVMPVANPPFKERALPCDS 234
Qy 142 -----LEPL----- 145
Db 235 ARPVEYSRPTLEVSPNVCHSNYSPKETIPBEARSDMHVSYAELGLKPAAPARNAPYF 294
Qy 146 -----BAEPT----- 177
Db 295 PCDKASKEERPSSEDETAHFEPNAPLNKGLVSPQSKDCQNPSTEACSSKNAC 354
Qy 178 -----SQGPPSPASDPKCNWKYKIVLNS-----QASQAGSLVGERSSGQPCP 223
Db 355 ILQSGSPPAKSPDPKACSKWKYKIVLNSLNQNAKPGGPEQAEGLGRLSPRAYTAPPAC 414
Qy 224 QARLPBGDEASSSSSSSSSEEGPIGPOSRL-----SPTAATVQF----- 266
Db 415 QPMW-EPENLDLQSPTKLSAGEDSTIP-QASRLNINVRNMTGSPRSSSESHSPLYMHP 472
Qy 267 ---KCG-APASTPYLLTSQ---AODTSGSPGERARPLPGSFFSCONCEAVAGCSSG 316
Db 473 PKTSCGQSQAEMCLHTAGPTFAEMGTQSEYSDSSCENGAFFCNEDCFSEAS 532
Qy 317 LD-SLVPGDEKPYKCOLCRSSFRYKGNLASHRTVHTGKPYHCSICGARNRPNALKTH 375
Db 533 LKRHTLQTHSDKPYKDCRQASFRYKGNLASHRTVHTGKPYHCSICGARNRPNALKTH 592
Qy 376 SRIHSGKPYKCTCGSRFVQVRSOPPSGPOGKPARGVGQKGFCSQRODLKSPSOV 435
Db 593 TRIHSGKPYKCTCGARFV-----QV 614
Qy 436 AHLRAHLVHTGKPYKCTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQLR 495
Db 615 AHLRAHLVHTGKPYKCEICGTRFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQLR 674
Qy 496 LHLRQKGAATNTKVHYHI 514
Db 675 LHLRQKGAATNTKVQYRV 693
RESULT 9
US-10-441-854-9
; Sequence 9, Application US/10441854
; Publication No. US20040028607A1
; GENERAL INFORMATION:
; APPLICANT: Brian J. No. US20040028607A1th
; APPLICANT: Eric M. Verdin
; TITLE OF INVENTION: Methods of Modulating Tubulin
; FILE REFERENCE: Deacetylase Activity
; CURRENT APPLICATION NUMBER: US/10/441,854
; CURRENT FILING DATE: 2003-05-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 803
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-441-854-9
Query Match 20.6%; Score 575.5; DB 15; Length 803;
Best Local Similarity 29.4%; Pred. No. 6.4e-27;
Matches 180; Conservative 64; Mismatches 205; Indels 163; Gaps 20;
Qy 16 EFTRHSSDVLGNLNLRLRGLITDVTLLVGGQPLRAHRAKAVLIACSGFFYSIFRGRAG-VG 74

Db 2 DFFQHSOHVLEQNQORQLGLLDCCTFVVDGVHFKAKAVLAACSEYFKMLFVDQKDVH 61
Qy 75 VDVLSPGPGPEARGFAPLLDFFMYTSRLSPATAPVLAATAATYLOMEHVYVQACHRFIOAS 134
Db 62 LDI-----SNAAGLQGMLEFMYTAKLSLSPENVDDVLAVALTFLQMDIITACH----- 109
Qy 135 YEPLGISLPLEAPPTPP-----TAPPPGSPRRSEGHDPPTESRSCSGPPSPASPD 188
Db 110 -----ALKSL-AEPATSPGNAEALATEGGDKRAKEKVATSTLRLQAGASTPIGP- 161
Qy 189 PKACNKKYKIVLNSQASQAGSLVGERSSGQPCQARLP-----SGDEASSSSSSSSSSSS 245
Db 162 --SRDLKEERGQQAASAGAEQTEKADAPREPPELVKDPDPTSGMAAAEAALSESSE 219
Qy 246 EEGPI-PGQPSRLSPTAATVQFKGA-PAST-----PYLLTSQA-----ODTSGSPS--- 290
Db 220 QEMEVEPARKGEBEQEQQEGBEGAGPAEVKEEGSQLENGEAPENENESAGTDSGOE 279
Qy 291 --BRAPLP-----GSEF-----FSCQNC 307
Db 280 LGSEARGLSGTGYDRTESKAYGVTHKCEDCGKEFTHTGNFKRHIRHTGKPFSCREC 339
Qy 308 EAV-----AGCSSGLDSLVPGEDEKPYKCOLCRSSFRY----- 340
Db 340 SKAFSPAAKAEKHTSP---LKPGCEGCKSYRLISLLNLKXKHSGEARYCEDCG 396
Qy 341 -----KGNLASHRTVHTGKPYHCSICG-----ARFN 367
Db 397 KLFTTSGNLKRLQVHSGKPYKDYCGRSFSDPTSKMRHLETHDTHKEHKCPHCKKFN 456
Qy 368 RPNALKTHSRIHSGKPYKCTCGSRFVQVRSOPPSGPOGKPARGVGQKGFCSQROD 427
Db 457 QVGNLKAHLKIHTADGLKRCGCKQFT-----TSGNLKRLHRIHSGKPYVICIHCQRQ 510
Qy 428 LKSPSQVAHLRAHLVHTGKPYKCTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLH 487
Db 511 FADPGA-----LQHRVHTGKPYKCCQCMCKAFTQASSLIAHVROHTGKPYVCERCGR 566
Qy 488 FRHKSQRLHLR 499
Db 567 FVQSSQLANHR 578
RESULT 10
US-10-486-977-12
; Sequence 12, Application US/10486977
; Publication No. US20050123912A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; BAROSSO, Ines;
; APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;
; APPLICANT: BLAKE, Julie J.; BOROWSKY, Mark L.;
; APPLICANT: BURFORD, Neil; DUGGAN, Brendan M.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: FORSYTHE, Ian J.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann E.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; HONCHELL, Cynthia D.;
; APPLICANT: ISON, Craig H.; KHAN, Farrah A.;
; APPLICANT: LAL, Preeti G.; LEE, Ernestine A.;
; APPLICANT: LEE, Sally; LEE, Soo Yeun;
; APPLICANT: LI, Joana X.; LU, Dyung Aina M.;
; APPLICANT: LU, Yan; LEHR-MASON, Patricia M.;
; APPLICANT: NGUYEN, Daniel B.; RAMKUMAR, Jayalaxmi;
; APPLICANT: SPRAGUE, William W.; TANG, Y. Tom;
; APPLICANT: THANGAVELU, Kavitha; THORNTON, Michael B.;
; APPLICANT: TRAN, Uyen K.; CHAWLA, Narinder K.;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YUE, Monique G.; YUE, Henry;
; TITLE OF INVENTION: NUCLEIC-ACID ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1146 USN
; CURRENT APPLICATION NUMBER: US/10/486,977
; CURRENT FILING DATE: 2004-02-17

PRIOR APPLICATION NUMBER: PCT/US02/25829
PRIOR FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: US 60/313,111
PRIOR FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: US 60/314,682
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/314,756
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/315,105
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/316,751
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/316,856
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328,185
PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PERL Program
SEQ ID NO 12
LENGTH: 765
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 3217430CD1

US-10-486-977-12
Query Match 19.5%; Score 543.5; DB 18; Length 765;
Best Local Similarity 28.7%; Pred. No. 5.5e-25;
Matches 170; Conservative 67; Mismatches 22; Indels 131; Gaps 19;

QY 20 HSSDVLNLELRGLITDVTLLVGGQPLRAHKAVALIACGFFVSIPIRGAGVGVVLS 79
DB 7 HSCLLQOLHEQRIGLQGLCCDWVKGVCFAHKNVLAAPFQYRSFQNSSQKNDVFH 66
QY 80 LPGGPEARGFAPLDFMYTSRLRSLSPATAPAVLAATYLOMEHVVOACHRFIQAS--YEP 137
DB 67 L-DVKNVSGIQLDFMYTSHLDNQDNIQVMLDTAQCLOVNVLSLCHTEKLSATVQP 125
QY 138 LG-----ISURLEAEPT-----PP-----TAPPPSPRSEHPD--PTESRSC 177
DB 126 PGMPCNSTLSLQSTLTPTDTCVISENYPHLLQECSDAQKNTLDESHFASPSVNRHH 185
QY 178 SQGPPSPASDPK--AC-----NWKYKVIIVNSOASO-----AGSLVGRSSGQPCPOAR 226
DB 186 SAGEISKQAPDTSQSGCTELPFKQPNYYIKLNFYSKYQYHKAAGPSQERVVEQPF--AF 243
QY 227 LPSGDEASSSSSSSSSSSE---EGP-----IPGQSRSLPTAATVQPK- 267
DB 244 STSDLTTVESQPCAVSHSECTLESPEHLPSNFLAQPVNDSAPHESDATCQPFKQWRL 303
QY 268 -----CGAPASTPYL---LTSQAQDT--SGSPSERARPLPGSEFFS 303
DB 304 KKAHLKXLFKQKAEQVSEPKSDGLTKRLESASKNTLEKASSQSAEKESEEVVS 363
QY 304 CONCAVAGCCSGLSLVPGDE-----DKPKCOLCRSSPKYKGNLASHRTVHTGKPY 357
DB 364 CENFCISETERPEDPAALDQSQTLQSQRYACELCGKPFKHPNLELKRSHTEGKPF 423
QY 358 HCSICGARFNRPANLKTSHRHSIGKPKYKCEGSRFVQVRSQPPGQKPARGV--- 414
DB 424 ECNIGKHFQAGNLQTHLRHSIGKPKYKCEICGKFA-----ASGDVQRH 469
QY 415 -----GQKGGFCSSQKDLKSPPSQVAHLRAHV-----LIH 445
DB 470 IHHGSEKPHLCDICRGFSNFSNLKEHKKHTADKVFCTDCGKSFNNQKLVKRRIRH 529
QY 446 TGEKYPYPCCTTRFRHLQTLKSHVRIHTGKPYKCDPCGLFRHKSQRLRH 497
DB 530 TGERPYSCSACGKFCGGGDLRRHVRHTGKPYKCEICNCKNCTFRSAVLRH 581

RESULT 11

US-10-965-898-4
Sequence 4, Application US/10965898
Publication No. US20050084936A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
Bandman, Olga
Hillman, Jennifer L.
Au-Young, Janice
Tang, Y. Tom
Yue, Henry
Shah, Purvi
Guegler, Karl J.
Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY PROTEINS
NUMBER OF SEQUENCES: 150
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/965,898
FILING DATE: 18-Oct-2004
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/001,403
FILING DATE: 31-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0455 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT01
CLONE: 609911
SEQUENCE DESCRIPTION: SEQ ID NO: 4 :
US-10-965-898-4

Query Match 18.3%; Score 509.5; DB 17; Length 539;
Best Local Similarity 28.3%; Pred. No. 4.5e-23;
Matches 160; Conservative 53; Mismatches 190; Indels 163; Gaps 18;
QY 1 MGSPAAPEGALGVYREPTRHSSDVLNLELRGLITDVTLLVGGQPLRAHKAVALIACS 60
DB 1 MGSP--BDDLIGI--PPFHSSELLSCLEQRLGHLCDLTIRTOGLEVYRTHRAVLAACS 56
QY 61 GFFYSIF-----RGRAGVGVVLSLPGGPEARGFAPLDFMYTSRLRL 103
DB 57 HYFKKLFTEGGGAVMGAGSGGTATGAGAGVCELDVF--GPEALG--ALLEFAYTATLTT 113
QY 104 SPATAPAVLAATYLOMEHVVOACHRFIQASYEPLGISLRLEAEPPTP----- 152
DB 114 SSANPAVLAQARLLEIPCVAACMEILQGS-----GL-----EAPSFDDDDCERARQY 162
QY 153 -----PTAPPPGSPRRSGHPDPPTESRSCSGPPSPASDPKACNWKYKVIIVNSQASQ 208
DB 163 LEAFATATASGVNGEDSPQVPLP-----PPPPPPPRPVARRSRKPRKAFLOTKGAR 215


```
Db      256 GD-----YRCPDQSPDRVGTMEQVSKNEGCOAGAELELSKAGPEEEEEDEERGE 311
Qy      254 QSRLSPTAATVQPKC-----GAPASTPYLLTQAOQDTSGSP--- 289
Db      312 KKSNN-----FKSICEKAFLEYEKSFLKHSKXHHGVATEVVYRC-----DTCQOTFAN 359
Qy      290 ----SERARPLPGSE--FFSCQNC-----EAVAGCSSGLDSLVPGEDDKPKYCQL 333
Db      360 RCNLKSHQSHVHSSERHPFCELCGKKFKKDKVGRHVLQVHEG-----GGE--RHRGQ 411
Qy      334 CRSSFRYKGNLASHRTVHTGEKYHCSICGAPFNPANLKTHSRHSGEKPKYCETCGSR 393
Db      412 CGKGLSSKTALRLHERHTHTGDRPYGCTECGARFQPSALKTMRIHTGEKPFVCECGAR 471
Qy      394 FVQVRSQPSPGFGKPARGGVGKGFGCS-----SORODLK-----SPP----- 432
Db      472 FTQ-----NHMLIYHKRCHTGERPFMCETCGKSPASKEYLKHNRHHTGSKPFKEVCF 525
Qy      433 ----SQVAHLRAHVLHTGEKPYPCPTCGTRFRHLQTLKSHVRIHTGEKPYHCDPCGLHFR 489
Db      526 RTFAQRNSLYQHIKVHTGERPYCCDQCGKQFTQNALQHRRIHTGERPFMCNACGRFTT 585
Qy      490 HKSQRLRLRLRQKHGAATNTKVH 511
Db      586 DKSTLRRH-----TSIH 597
```

Search completed: August 9, 2005, 14:09:40
Job time : 164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 13:56:27 ; Search time 42 Seconds
(without alignments)
1186.673 Million cell updates/sec

Title: US-09-815-379-4
Perfect score: 2789
Sequence: 1 MGSPAPEGALGVYREFTRH.....RQKHGAATNTKVHYHILGGP 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2556	91.6	480	2 JC7812	BCL6 homologous zi
2	1128.5	40.5	706	2 I32586	B-cell CLL/lymphom
3	1122.5	40.2	706	2 A48752	B-cell CLL/lymphom
4	559	20.0	794	2 S59069	Z13 protein - mous
5	482	17.3	610	2 JC7315	myoneurin - human
6	481	17.2	688	2 A56360	zinc finger protei
7	434.5	15.6	614	2 JH0500	zinc finger protei
8	432.5	15.5	555	2 I53869	zinc finger protei
9	424	15.2	595	2 G2075	transcription repr
10	421.5	15.1	673	2 S36336	probable transcrip
11	419	15.0	676	2 I50643	gammaPBP-C - chick
12	417.5	15.0	701	2 T14757	hypothetical prote
13	415	14.9	803	2 S26823	zinc finger protei
14	413.5	14.8	428	2 A32891	finger protein 1,
15	412.5	14.8	693	2 I37570	zinc finger protei
16	408	14.6	728	2 A48830	probable transcrip
17	407	14.6	209	2 S47068	finger protein HZF
18	404	14.5	668	2 T08725	probable finger pr
19	401	14.4	1191	2 S53305	zinc finger protei
20	399	14.3	469	2 I38600	zinc finger protei
21	399	14.3	485	2 A40751	finger protein MZF
22	397	14.2	543	2 B34612	zinc finger protei
23	396.5	14.2	613	2 JC7992	negatively regulat
24	396	14.2	449	2 S41647	zinc finger 5 prot
25	394.5	14.1	594	2 T12488	hypothetical prote
26	394	14.1	196	2 S06561	finger protein (cl
27	393.5	14.1	466	1 TWFF	transcription fact
28	393.5	14.1	475	2 S03679	finger protein (cl
29	392	14.1	292	2 S43826	finger protein OZF

ALIGNMENTS

RESULT 1

JC7812

BCL6 homologous zinc finger protein BAZF - human

C;Species: Homo sapiens (man)

C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004

C;Accession: JC7812

R;Sakashita, C.; Fukuda, T.; Okabe, S.; Kobayashi, H.; Hirose, S.; Tokuhisa, T.; Miyase

Biochem. Biophys. Res. Commun. 291, 567-573, 2002

A;Title: Cloning and characterization of the human BAZF gene, a homologue of the BCL6 onc

A;Reference number: JC7812; PMID:11855826; MUID:21845446

A;Accession: JC7812

A;Molecule type: DNA

A;Residues: 1-480 <SAK>

A;Cross-references: UNIPROT:O8N143; DDBJ:AB076580; DDBJ:AB076581

C;Comment: This protein, a zinc finger protein containing a conserved amino terminal B7B, rly in the induction of megakaryocytic differentiation to produce mature platelets. This

cell fate.

C;Genetics:

A;Gene: bazf

A;Map position: 17p13.1

A;Introns: 60/2; 134/2; 256/2; 297/3; 353/1; 396/3; 442/3

C;Keywords: differentiation

Query Match 91.6%; Score 2556; DB 2; Length 480;

Best Local Similarity 92.7%; Pred. No. 4.9e-136;

Matches 480; Conservative 0; Mismatches 0; Indels 38; Gaps 1;

Qy 1 MGSPAPEGALGVYREFTRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACS 60

Db 1 MGSPAPEGALGVYREFTRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACS 60

Qy 61 GFFYSIFRGAGVGVVDVLSLPGGPEARGFAPLDPMYTSRLRLSPATAPAVLAATYIQM 120

Db 61 GFFYSIFRGAGVGVVDVLSLPGGPEARGFAPLDPMYTSRLRLSPATAPAVLAATYIQM 120

Qy 121 EHVVQACHRTIOASYEPLGISLRPLEAEPTPTAPPSPRRSEGHDPPTESRSCSQG 180

Db 121 EHVVQACHRTIOASYEPLGISLRPLEAEPTPTAPPSPRRSEGHDPPTESRSCSQG 180

Qy 181 PPSPASDPKACNWKYKYLNSQAQSLVGRSSGQPCQARLPSPGDEASSSSSSS 240

Db 181 PPSPASDPKACNWKYKYLNSQAQSLVGRSSGQPCQARLPSPGDEASSSSSSS 240

Qy 241 SSSSSEGGPIPGQSRISPTAATVQFKCGAPASTPYLLTQAQDTSQSPSRAPLPQSE 300

Db 241 SSSSSEGGPIPGQSRISPTAATVQFKCGAPASTPYLLTQAQDTSQSPSRAPLPQSE 300

Qy 301 FFSQNCCEAVAGCGSLDLSVPGDEDPKPKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360

Db 301 FFSQNCCEAVAGCGSLDLSVPGDEDPKPKQLCRSSFRYKGNLASHRTVHTGKPYHCS 360

Qy 361 ICGARFNRPNALKTHSRHSKEPKYKCTCGSRFVQVRSPQSPFGQKPARGGVQKGGF 420

Db 361 ICGARFNRPNALKTHSRHSKEPKYKCTCGSRFVQVRSPQSPFGQKPARGGVQKGGF 420

QY 61 GFFYSIFRGRAGVGVVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATAATYLOM 120
 DB 55 GLFYSIFPDQLKCNLSVNLQPEINPEGFCILLDFMYTSRLRLREGNIMAVMATAYLOM 114
 QY 121 EHVVOACHRFQAS 134
 DB 115 EHVVDTRCKFKIKASEAEAMVSAIKPPREFFLNSRLMPOQIMAYRGREVENNLPLRSAGP 174
 QY 135 -----YELGIS----- 141
 DB 175 CESRAPAPSLYGLSTPPASYSMYSHLPVSSLLFSDDEFRDVRMPVANPPKERALPCDS 234
 QY 142 -----LRPL----- 145
 DB 235 ARPVPGEYSRPTLEVPNVCHSNISPKETIPERASDMHYSVAEGLKPAAPASARNAPYF 294
 QY 146 -----EAPPT-----PQSPRRSEGHDPPTESRS----- 176
 DB 295 PCDKASKEEERFSEDEITAHFEPNAPLNKGLVSPQSPQKDCQPNSTPEACSSKNVAC 354
 QY 177 ---CSQGPSPASPDPAKCNWKKYKIVLNS-----QASQAGSLVGERSSGQPCP 223
 DB 355 ILQASGSPPAKSPDPAKCNWKKYKIVLNSQNAKPGGPEQAEGLRLSPRAYTAPPAC 414
 QY 224 QARLPBGDEASSSSSSSSSSSESGPIPGOSRL-----SPTAATVQF----- 266
 DB 415 QPM-EPENLDLQPTKLSAGEDSTIP-QASRLNNIVNRMTGSPRSSSESHSPLYMHP 472
 QY 267 ---KCG--APASTPYLLTSQ---AOTSGSPGERARPLPGSFFSCQCEAVAGCSG 316
 DB 473 PKTSCGSGSPQHAEMCLHTAGTFAERMGETQSEYSDSSCENGAFPCNECDRFBSEAS 532
 QY 317 LD-SLVPDEDEPKYKCOLCRSEFPYKGNLASHRTVHTGKPYHCSI CGARENRPANLKT 375
 DB 533 LKRHTLOTHSDPKYKDCQASFRYKGNLASHRTVHTGKPYHCSI CGAQRNRPANLKT 592
 QY 376 SRIHSGEKPYPKCTCGSRFVQVRSPPSGFGKARGVGQGGFCSSQRODLKSPPSQV 435
 DB 593 TRIHSGEKPYPKCTCGARFV-----QV 614
 QY 436 AHLRAHVLHTGKPYPCPTCTGFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQLR 495
 DB 615 AHLRAHVLHTGKPYPCPTCTGFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQLR 674
 QY 496 LHLRQKGAATNKKVHYH 514
 DB 675 LHLRQKGAATNKKVHYH 693

RESULT 4
 S59069
 213 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S59069
 R:Schulz, T.C.; Hopwood, B.; Rathjen, P.D.; Wells, J.R.E.
 Biochem. J. 311, 219-224, 1995
 A:Title: An unusual arrangement of 13 zinc fingers in the vertebrate gene Z13.
 A:Reference number: S59069; PMID:7575457
 A:Accession: S59069
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-794 <SCH>
 A:Cross-references: UNIPROT:Q60821; GB:U14556; NID:G608136; PIDN:AAA85493.1; PID:G608137
 F:10-108/Domain: POZ domain homology <POZ>

Query Match 20.0%; Score 559; DB 2; Length 794;
 Best Local Similarity 26.6%; Pred. No. 3.8e-24;
 Matches 164; Conservative 73; Mismatches 198; Indels 182; Gaps 17;
 QY 16 EFTRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACSGFFYSIFRGAAG-VG 74
 DB 2 DFPQHSQVRLBQNLQRLGLLDCDCTFVVDGVDFKAHKAVALAACSEYFKMLFVDQKVH 61

QY 75 VDVLSPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATAATYLOMEHVVOACHRFQAS 134
 DB 62 LDI-----SNAAGLGQVLEFMYTAKLSLSPENVDDVLAVASFLQMDIVTACHT-LKSL 114
 QY 135 YELGLISLRLPLEAPPTPTAPPSPRRSEGHDPPT----- 172
 DB 115 AEPSTTTGESADA-----SAVEGGDKRAKDEKAAATMLSRIGQARGSSSTGPGRELKE 167
 QY 173 -----ESRSCSQ-----PSPASPPKACNWKYKIVLNSQASQASGL----- 212
 DB 168 ERGQASASAGABQTEKADAPREPPVELKPDPTS-----SMAAAEAEALSES 216
 QY 213 -----VGRSSQPCPQ-----ARLPBGDEASSSSSSSSSESGPIPGP 253
 DB 217 SEQEMEVEPASKGEDQGEGBGAGPATVKERGMHLDNGEPEENESAGTDSGQELGMEGQ 276
 QY 254 QSRLSPTAATVQPK-----CGAPASTPYLLTSQAQDTSGSPSERARPLPGSEFF 302
 DB 277 NLRSGTVGDRTEKAYGSIHKCEDCGKFT-----HTGNFKRHIRIHTGSKPF 325
 QY 303 SCQCEAV-----AGCSSGLDSLVPGEDPKYKCOLCRSSPRY----- 340
 DB 326 SCRECKAFSDPAACKAHEKTHSP-----LKPYGCECKSYRLISLLNLHKKHSGEAYR 382
 QY 341 -----KGNLASHRTVHTGKPYHCSI CG----- 363
 DB 383 CGDCGKLTFTSGNLKRHLVHSGOKPYQCDYCGRSFSDPTS KMRHLETHDTHDKHKCPHC 442
 QY 364 -ARPNRPANLKTSHRHSGEKPKYKCTCGSRFVQVRSPPSGFGKARGVGQGGFCPS 422
 DB 443 DKXFNQVGNLKAHLKIHADGPKLCKECGKQFT-----TSGNLKRHLRHSGEKPYVCT 496
 QY 423 SQODLKSPPSOVAHLRAHVLHTGKPYPCPTCTGFRHLQTLKSHVRIHTGKPYHCD 482
 DB 497 HCQRQFADPGA-----LQRHVRITHTGKPCQCVICGKAFQTOASSLIAHVRQHTGKPYCE 552
 QY 483 PCGLHFRHKSQLRHLR 499
 DB 553 RCGKRFVQSSQLANHIR 569

RESULT 5
 JC7315
 myoneurin - human
 C:Species: Homo sapiens (man)
 C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
 C:Accession: JC7315
 R:Alliel, P.M.; Seddiqui, N.; Goudou, D.; Cifuentes-Diaz, C.; Romero, N.; Velasco, E.; Ri
 Biochem. Biophys. Res. Commun. 273, 385-391, 2000.
 A:Title: Myoneurin, a novel member of the BTB/POZ-zinc finger family highly expressed in
 A:Reference number: JC7315
 A:Accession: JC7315
 A:Molecule type: mRNA
 A:Residues: 1-610 <ALL>
 A:Cross-references: UNIPROT:Q9NPC7; GB:AF14884
 C:Comment: This protein, belonging to the family of eukaryotic BTB/POZ and zinc finger p
 and a regulator in the genital tract.
 C:Keywords: tandem repeat; transcription regulation; zinc finger

Query Match 17.3%; Score 482; DB 2; Length 610;
 Best Local Similarity 26.8%; Pred. No. 5.9e-20;
 Matches 149; Conservative 65; Mismatches 212; Indels 130; Gaps 16;
 QY 16 EFTRHSSDVLGNLNLRLRGILTDVTLVGGQPLRAHKAVALIACSGFFYSIFRGAAG-VG 73
 DB 2 QYSHHCEHLRLNKRQEAFLCDCTTIVGEGFQKARNVNLASFSEYFGALYRSTSENNV 61
 QY 74 GVDVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPAVLAATAATYLOMEHVVOACH----- 128
 DB 62 FLDSQV-----KADGFKLLEFIYTGTLNLDNSWNVKEIHQAADYLVKEEVVTKCKMKED 117
 QY 129 -----RFIQASYEPLGISLRP-----LEAEPPTPTAPPGP 159

```
Db 118 FAFIANPSTBTSSITGNIELNQOTCLLTLRDYNNRREKSEVSTDLIQANPKQALAKKSS 177
Qy 160 -SPRRSEGHDPPTESRSCQGP-----PSPASPDPPKACNWK---KY 197
Db 178 QTKKKKAFNSPKTGKNTQVQPSDILENASVELFDANKLTPVVEQVAQINDNSELEL 237
Qy 198 KYIVLNSQASQ--AGSLVGERSSGQPCQARLPS---GDEASSSSSSSSSSSEEGPIGP 252
Db 238 TSWENTFPQAQDIVHTVTVKRRKGSQPCALKEHSMNSIAKSPYEAENSCEE--LDQ 295
Qy 253 POSRLSPATAATVQKCGAPASTFYLLTSAQDTSGPSERARP-----LPGSEFFSCQCE 308
Db 296 RYSKAKPMCMT---CGKVFSEASSLRHMRTHKG-----VKPYVCHLCGKAFTQCNOLK 346
Qy 309 AVAGCSSGLDVLPGDEDPKPKQCRSSFRYKGNLASHRTVHTG-EKPYHCISICGARFN 367
Db 347 THVRTHG-----EKPYKELCDKGFAQKQCVLPHSRMHGEEAPYKCDVCNLOFA 397
Qy 368 RPANLKTSHRIHSGEKPKYKCTCGSRFVQVRSQPSGFGKPARGGVQKGFCSQROD 427
Db 398 TSNLKIHAHSGEKPKYVCDRCGRF----- 424
Qy 428 LKSPSOVAHLRAHVLHTGKPYCPTCGTRFRLHTLTKSHVRIHTGKPVHCDPCGLH 487
Db 425 -----AQASTLYHVRRHTGKPYVCDTCGKAFVSSSLITHSRKHTGKPYICIGCKS 479
Qy 488 FRHKSQLRLHLROKHG 503
Db 480 FISSGELNKHFRSHTG 495

RESULT 6
A56360
zinc finger protein (clone 18) - human
N:Alternate names: GLI-related finger protein HKR3
C:Species: Homo sapiens (man)
C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A56360; E31201
R:Sugawara, M.; Scholl, T.; Ponath, P.D.; Strominger, J.L.
Mol. Cell. Biol. 14, 8438-8450, 1994
A:Title: A factor that regulates the class II major histocompatibility complex gene DPA
A:Reference number: A56360; MUID:95059073; PMID:7969177
A:Accession: A56360
A:Molecule type: mRNA
A:Residues: 1-688 <SUG>
A:Cross-references: UNIPROT:P10074; GB:L116896; NID:G292934; PIDN:AAAG5124.1; PID:G292935
R:Ruppert, J.M.; Kinzler, K.W.; Wong, A.J.; Bigner, S.H.; Kao, F.T.; Law, M.L.; Seunanez,
Mol. Cell. Biol. 8, 3104-3113, 1988
A:Title: The Gli-Kruppel family of human genes.
A:Reference number: A93103; MUID:89096896; PMID:2850480
A:Accession: E31201
A:Molecule type: DNA
A:Residues: 461-516 <RUP>
C:Genetics:
A:Introns: 506/1
A>Note: the list of introns may be incomplete
C:Keywords: DNA binding; zinc finger
F:12-111/Domain: POZ domain homology <POZ>

Query Match 17.2%; Score 481; DB 2; Length 688;
Best Local Similarity 27.7%; Pred. No. 7.6e-20;
Matches 160; Conservative 72; Mismatches 216; Indels 130; Gaps 22;

Qy 17 FTRHSDVLGNELRLGILTDVLLVCGQPLRAHKAVLIACSGFFYSIFRGRAGVGD 76
Db 5 FVQHSVRVLQELNKKREKQYCDATLDVGLGVFKAHWSVLACCSHFQSLYGDGSGSV- 63
Qy 77 VLSLPGGPARGFAPLLDDPMYTSRLRLSPATAPAVLAATYLOMEHVWQACHRF----- 131
Db 64 --VLPAG-FAEIFGLLLDFPYTHGLALTSNGRDVLLAARELVRPEAVELCOSFPKTSV 120
Qy 132 -QAS-----YEPLG-----ISLRPLEARF--PTPTAPPGPS 161
```

```
Db 121 GQAGGQSGGLGPPASQNVNHHVKEPAGLEBEEVSTGLVPRDQEPGRGSHSPQRQLHSP 180
Qy 162 RRSEGHDPPTTE-----SRSCS-----QGPPSPASPDPPKACNWKYKIVL-- 202
Db 181 AQSEG-----PSSLGKGLKQALKPCSLDKKPEDCKVPPRPLEAEGAQLOGGSNWEVVVQ 236
Qy 203 -----NSQASQAGSLVGERSSG---QPCQARLPSGDEASSSSSSSSSEEGPIGP 253
Db 237 VEDDGDGYSYNSEPEAVLTRKSNVIRKPC-----AAEPALSAGSLAAE-----P 280
Qy 254 QSRLSPATAATVQKCGAPA-----STPYLLTSAQDTSGPSERARPLPGSEFFSCQNC 307
Db 281 AENKGTAVPVE-----CPTCHKFLSKYLLKVNHRKHTGKPF--CPKCKCIVFRKNL 334
Qy 308 --EAVAGCSSGLDVLPGDEDPKPKQCRSSFRYKGNLASHRTVHTGKPYHCISICGAR 365
Db 335 LEHEARNCMNRSEQVLP-----CSVCQBTFFRRMELRVHMYSHTSHTGEMPYKSCSSCQ 386
Qy 366 PNRPANLKTSH--RIHSGEKPKYKCTCGSRFV-QVRSQPSGFGKPARGGVQKGFCS 423
Db 387 FMQKDLQSHMKILHGAQKPHACPTCAKCFLSRTELQHEAFKRGKELFVCEEGHRS 446
Qy 424 QRODLK-----SPP-----SOVAHLRAHVLHTGKPYCPTCGTRFRLH 464
Db 447 SRNGLQWHIIXAKHNRNPHVCEPCSHAFTQKANLNMHLRHTHTGKPFQCHLGGKTRTQA 506
Qy 465 TLKSHVRIHTGKPYHCDPCGLHFRHKSQRLRLHRLQKH 502
Db 507 SLDKHNRHTHTGERPFCSECFCEQRTKGLLRHVASRH 544

RESULT 7
JH0500
zinc finger protein 29 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: JH0500
R:Denny, P.; Ashworth, A.
Gene 106, 221-227, 1991
A:Title: A zinc finger protein-encoding gene expressed in the post-meiotic phase of sperm
A:Reference number: JH0500; MUID:92039080; PMID:1937051
A:Accession: JH0500
A:Molecule type: mRNA
A:Residues: 1-614 <DEN>
A:Cross-references: UNIPROT:Q07230; GB:X55126; NID:G55470; PIDN:CAA38920.1; PID:G55471
A:Experimental source: Testis
C:Keywords: DNA binding; zinc finger
F:217-244/Region: zinc finger
F:245-272/Region: zinc finger
F:273-300/Region: zinc finger
F:301-328/Region: zinc finger
F:329-356/Region: zinc finger
F:357-384/Region: zinc finger
F:385-412/Region: zinc finger
F:413-440/Region: zinc finger
F:441-468/Region: zinc finger
F:469-496/Region: zinc finger
F:497-524/Region: zinc finger
F:525-552/Region: zinc finger
F:553-580/Region: zinc finger
F:581-608/Region: zinc finger

Query Match 15.6%; Score 434.5; DB 2; Length 614;
Best Local Similarity 30.3%; Pred. No. 2.7e-17;
Matches 122; Conservative 42; Mismatches 133; Indels 105; Gaps 15;

Qy 162 RRSEGHDPPTESRSCQGPSPASP--DPKACNWKYKIVLNSQASQAGSLVGERSS- 218
Db 170 QQSDQSDS---FERDCSGGAQGHAGEDPR-----VVPSEGEVGLIGLQGT 216
Qy 219 -GQ---PCQA-----RLPSG-----DEASSSSSSSSSEEGPIGPQ 254
```

Db 217 LGKPYECPCQCKTSRKSRLTHRTHTGKYYKCDCEGKSFSDGNSFSRHQTHHTGK 276
Qy 255 SRLSPATAATVQFK---CGAPAS--TPYLLTSQAQDTSGSPSERA-----R 294
Db 277 -----PYKCRDCKGKSFSSANLTHQRIHTGKPFQCAECGKSFSSPNLIAHOR 326
Qy 295 PLPGSEFFSCNCEAVAGSSGLDS---LVPGDEKPKYQCLCRSSFRYKGNLASHRTVH 351
Db 327 THTGKPYSCPECGKSFGRSSLNTHQIHTG--EKPYACKECGSEFSYNSNLIRHORIH 384
Qy 352 TGEKPYHCSICGARENRPANLTHRTSHSGKPKYKCTCGSRFVQVRQPPSGFGKAPAR 411
Db 395 TGEKPYKTECQKQKFSQSALITHRTHTGKPYOCGEGKGF-----427
Qy 412 GGVGKGKGFSSQRODLKSP-----SQVAHLRAHVLHTGKPYCPTCGTRFR 461
Db 428 ---SRSSNLATHRRTHLVEKPKYKGLGKSFSSQSSSLIAHQHTGKPYELCTGSEFS 484
Qy 462 HLQTLKSHVRHTGKPYKPCDGLHFRHKSQLRLHLRQKHG 503
Db 485 WSSNLIKHQHTGKPYKRCGDKGFGSQRSQSLVVHQHTHG 526
RESULT 8
153869
Query Match 15.5%; Score 432.5; DB 2; Length 555;
Best Local Similarity 28.6%; Pred. No. 3.2e-17;
Matches 135; Conservative 52; Mismatches 178; Indels 107; Gaps 18;
C:Species: Mus musculus (house mouse)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: I53869
R:Brady, J.P.; Piatigorsky, J.
Gene 149, 299-304, 1994
A:Title: A mouse cDNA encoding a protein with zinc-fingers and a KRAB domain shows simi
A:Reference number: I53869; MUID:95047492; PMID:7959006
A:Accession: I53869
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-555 <RES>
A:Cross-references: UNIPROT:Q62518; GB:L28167; NID:9758660; PIDN:AAA67545.1; PID:g758661
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
Qy 79 SLPGGPEARGPAPLDFWYTSRLSLSPATAPAVLAATYQWHEVYVQACHRFQIASVEPL 138
Db 112 SVNTPTQOGKAP-----CQSRPSPAPSVVDENYIK-SHTISV-----149
Qy 139 GISLRPLEAEPTPTAPPKGS-----PRRSEGHDP-----PP 171
Db 150 -----PIHSGRPPLTLRAQGTSLTERQTRKDQYHQGSHRDEAGQKRESHTSWIPRPR 204
Qy 172 TESRSC-SQGPPSPAS----PDKACNWKYKIVLNQASQAGSLVGRSSGQPCQARL 227
Db 205 SDHRVCKSGKRCPSDGHGRDGVVPPQNGIHTQRKSSQHREAVTDLSSSD--PHQVL 262
Qy 228 PSGDEASSSSSSSSS-----SSEEGPIPGQSRLSPTAATVQFKCAPAST 274
Db 263 DSGGRACACEGRCGSPCGTDCIHQGGRTGEERKEFGHQVCTPVKPYTCEQC-----317
Qy 275 PYLLTSQAQDTSGSPSERARPLPGSEFFSCNCEAVAGSSGLDS---LVPGDEKPKYC 331
Db 318 KSLLDVQHLMHVKVHTTERP-----YNCETCGSAFQASHLQHORLHTG--EKPFKC 369
Qy 332 QLCRSSFRYKGNLASHRTVHTGKPYHCSICGARENRPANLTHRTSHSGKPKYKCTCG 391
Db 370 DACGKSFSSSHLRSHQVHTGKPYKCEGKSFICSSNLYIHOVHTGKPYKVCVDCG 429
Qy 392 SRVQVRQSPSGFGKARGGV--GQKGGFCS--SQRODLKSPSPQVAHLRAHVLHTG 447
Db 430 KEFSR-----PSSLQ---AHQGIHTGKSYVCTMCKGKYTLNS-----NLQVHLRVHTG 475
Qy 448 EKPYPCTCGTFRHLQTLKSHVRHTGKPYKPCDGLHFRHKSQLRLHLR 499

Db 476 EKPYSQCDVCGKGFSSQLOSHQVHTGKPYKCDVCKSGFGRSNLIHHR 527
RESULT 9
G02075
Query Match 15.2%; Score 424; DB 2; Length 595;
Best Local Similarity 38.5%; Pred. No. 1e-16;
Matches 90; Conservative 26; Mismatches 92; Indels 26; Gaps 5;
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C:Accession: G02075
R:Poncelet, D.A.
submitted to the EMBL Data Library, September 1995
A:Reference number: G09169
A:Accession: G02075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-595 <PON>
A:Cross-references: UNIPROT:Q03923; EMBL:U35376; NID:g1017721; PIDN:AAA79179.1; PID:g101;
C:Genetics:
A:Gene: GDB:ZNF85
A:Cross-references: GDB:I32279
A:Map position: 19p12-19p12
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
Qy 267 KCGAPASTPYLLTSQAQDTSGSPSERARPLPGSEFFSCNCEAVAGSSGL--DSLVPG 323
Db 206 ECG-----KAFNWSSTLTQKRIHTGKPYKCEGKAFNQSNNLKKHKIHTG 254
Qy 324 DEDKPYKQCLCRSSFRYKGNLASHRTVHTGKPYHCSICGARENRPANLTHRTSHSGK 383
Db 255 --EKPYKCEGKAFNFSSTLTTHKIHTGKPYKCEGKAFNRSSTLTTHRKIHTGSK 312
Qy 384 PYKCTCGSRVQVRQSPSGFGKARGGVQKGGFSSQRODLKSPSPQVAHLRAHVL 443
Db 313 PYKCEGKAFKQSSNLTTHKII-----HTGKPYKCKGKAF--NOSAHLTTHV 362
Qy 444 IHTGKPYCPTCGTFRHLQTLKSHVRHTGKPYHCDPGLHFRHKSQLRLH 497
Db 363 IHTGKPYKCEGKAFNFSHSLTHKIHTGKPYKCEGKAFNKSSTLTGK 416
RESULT 10
S36336
Query Match 15.1%; Score 421.5; DB 2; Length 673;
Best Local Similarity 25.3%; Pred. No. 1.6e-16;
Matches 155; Conservative 63; Mismatches 236; Indels 159; Gaps 19;
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S36336; S31989
R:Chen, Z.; Brand, N.J.; Chen, A.; Chen, S.J.; Tong, J.H.; Wang, Z.Y.; Waxman, S.; Zelen
EMBO J. 12, 1161-1167, 1993
A:Title: Fusion between a novel Krueppel-like zinc finger gene and the retinoic acid rec
A:Reference number: S36336; MUID:9320216; PMID:8384553
A:Accession: S36336
A:Molecule type: mRNA
A:Residues: 1-673 <CHE>
A:Cross-references: UNIPROT:Q05516; EMBL:Z19002; NID:g38517; PIDN:CAA79489.1; PID:g38518
C:Genetics:
A:Gene: PLZF
C:Keywords: zinc finger
F:20-118/Domain: POZ domain homology <POZ>
Qy 20 HSSDVLGNLNLRLGILTDVTLVGGQPLRAHKAVALIACSGFFYSIF-RGRAGVGVDVL 78
Db 16 HPTGLLCKANQMRAGTLCDVIMVDSQFHAHRTVLACTSKMFIILFHRNSQHTLDFL 75
Qy 79 SLPGGPEARGPAPLDFWYTSRLSLSPATAPAVLAATYQWHEVYVQACHRF---IQASY 135

```
Db 76 S-----PRTFOQILEYATYATLOAKEDLDLLYAAEILEYEEQCKMLETIOASD 129
Qy 136 E-----
Db 130 DNDTEATWADGAEBEEDRKARYLKNIFISKHSSESGYASVAGOSLPGMPVQSPSVST 189
Qy 137 -----PLGISLRLEAPPTPTAPPSPRRSEGHPPPTTESRCS 178
Db 190 SFGLSAMSPTKAADVSLMTIGQSLQGLTLPAGPEEPTLAGGRHPGVAEVTMMQVD 249
Qy 179 QGPPSPAPDPKACNWKYKYIIVLNSQASQAGSLVGRSSGQPCQAR-----LPS 229
Db 250 E---VPSQDSPGAAB-----SSISGGMGDKVEERKEGPGTPTSSVITSARELHY 297
Qy 230 GDEASS-----SSSSSSSSSEEGPIQPOSRLSPTAATVQKCGAPASTPYLLTS--- 280
Db 298 GREESAQVPPPAEAGQAQPTGPEHPAPPEKHLGIYVLPNKHADAVLSMSSVTSGLH 357
Qy 281 --QAQDTSGSPSERARPLPGS-----EFTSCNCEAVAGCSGLSLVPGDEDKPYKCOLC 334
Db 358 VOPALAVSMDFTYGLLPQGPPIQRELFPS-KLGEAVGKMS--ESRTTIGEQ-----CSVC 409
Qy 335 RSSFRYKGNLASHRTVHTGKPYHCSI CGARNRPNANKTHSRHIS-GEKPYKCTCGSR 393
Db 410 GVPELDNEAVEQHRKXLSHGMKTYGCELCGRFLDSLRMLMHLAHSAGAKAFVCDQCAQ 469
Qy 394 F-----VVRSPSPSG-----FOGKPA-----RGVGOKGFGCSSQRODLKS 430
Db 470 FSKEDALSTHRTQTHGTDMAVFCLLCGRFQAQSLQOHMEVHAGVRSYIISCEKNTFP-- 527
Qy 431 PPSQVAHLRAHVLHTGKPYCPCTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLFRH 490
Db 528 -PSHTA-LKRHLRSHTDHPYECFCGSCFRDESLKSHKRIHTGKPYECNGCDKXESL 585
Qy 491 KSQRLHLRQKHG 503
Db 586 KQLETHYRVHTG 598

RESULT 11
150643
gammaFBP-C - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence revision 13-Sep-1996 #text change 09-Jul-2004
C:Accession: I50643; I50641; I50642; S44264; S44265
R:Liu, Q.; Shalaby, F.; Puri, M.C.; Tang, S.; Breitman, M.L.
Dev. Biol. 165, 165-177, 1994
A:Title: Novel zinc finger proteins that interact with the mouse gamma F-crystallin from
A:Reference number: I50641; MUID:94374565; PMID:8088434
A:Accession: I50643
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-676 <LIU>
A:Cross-references: UNIPROT:Q90850; EMBL:X79011; NID:G475903; PIDN:CAA55644.1; PID:G4759
A:Accession: I50641
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 'MRVRELGLWA', 23, 'GSGR', 28-676 <LI2>
A:Cross-references: EMBL:X79051; NID:G479055; PIDN:CAA55653.1; PID:G479056
A:Accession: I50642
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 36-676 <LI3>
A:Cross-references: EMBL:X79050; NID:G479057; PIDN:CAA55652.1; PID:G479058
C:Superfamily: BRcore-2 protein; POZ domain homology
F:49-152/Domain: POZ domain homology <POZ>

Query Match 15.0%; Score 419; DB 2; Length 676;
Best Local Similarity 23.7%; Pred No. 2, 2e-16;
Matches 146; Conservative 67; Mismatches 184; Indels 218; Gaps 21;
Qy 9 GALGVREFTRHSSDVLGNLNLRLRGILTDVTLVAGGQPLRAHKAHLIACSGFFYSIFR 68
```

```
Db 34 GAMLEAMEVPSHQSLQLLQNTQRTKGLFCDVITVQNALFRAHKNILAAASAYLKSIV- 92
Qy 69 GRAGVGVDVLSLPGPEARG-FAPLLDFMYTSRL-RLSPA---TAPAVLAATYQLQMEHV 123
Db 93 ---VHONLLMLDHVMVSPGIFRULDFIYTRGLCEPGEQSLGALAAASYLQIQGL 148
Qy 124 VQACHR-----FIOASYE--PLGISLRPLE----- 146
Db 149 VALCKKXLRSGKYCHLRGGVAPYKLGRLRATTPVIOACYSCTPRPVDLQVEPAAPLN 208
Qy 147 -----A 147
Db 209 TQGELYASASQGTPLPHGLCPPEHSCPPCGLDLSSKSTGTPSAQLLPTDRLLPABPR 268
Qy 148 EPPTPP-----TAPPPGSPRRSEGHF---DPPTESRSCSQGPSP 184
Db 269 EPSLPPRRDPPVSGGLLAGHPAAYKDSPPGEP---GGPHATDTPFRSTPPCAE-PLP 324
Qy 185 ASDPKAC-NWKYK---YIVLNSQASQAGSLVGRSSGQP--CPQARLPS--GDEASS 235
Db 325 RGDGRELMYRMWKHEPLGPYL---DEGEAEKELEREKAEPPAAPQPRYSPVESNDLEP 381
Qy 236 SSSSSSSSSSEEGPIG-----POSRLSPTAATVQKCGAPASTPYLLTSQA 282
Db 382 DNSTSEETSGSEGPSGDALDRYCNHLGYEPESLGDNLVCI PCGKGFPSSEQLNAHVEA 441
Qy 283 QDTS---GSPSERARPLPGSEFFSCNCEAVAGCSGLSLVPGDEDKPYKCOLCRSSF 338
Db 442 HNEELVHKAAAEQAVP-----LDKGGAGL-----GDILRPYRCSCKSKY 483
Qy 339 RYKGNLASHRTVHTGKPYHCSI CGARNRPNANKTHSRHISGKPYKCTCGSRFVQR 398
Db 484 KDPATLRQHEKHTWLTRPYCTICGKFTQRTGTRHMRSHLGLKPFACDAGCMRF----- 539
Qy 399 SQPPSGFGQKPARGVGQKGGFCSSQRODLKSPSQVAHLRAHVLHTGKPYCPCTCGT 458
Db 540 -----TRQYRLTEHMRHISGKPYECQVCVG 565
Qy 459 RFRHLQTLKSHVRIH 473
Db 566 KFAQRNLISHMKWH 580

RESULT 12
T14757
hypothetical protein DKFZp572C163.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text change 09-Jul-2004
C:Accession: T14757
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: 218181
A:Accession: T14757
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-701 <WAM>
A:Cross-references: UNIPROT:Q9UG02; EMBL:AL110217
A:Experimental source: adult subthalamic nucleus; clone DKFZp572C163
C:Genetics:
A:Note: DKFZp572C163.1
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 15.0%; Score 417.5; DB 2; Length 701;
Best Local Similarity 41.6%; Pred. No. 2, 7e-16;
Matches 89; Conservative 23; Mismatches 77; Indels 25; Gaps 5;
Qy 298 GSEFFSCNCEAVAGCSGL---DSLVPGEDDKPYKCOLCRSSFRYKGNLASHRTVHTGE 354
Db 352 GEKPYECNECEKTFANHSALRAHQHNTG--EKLVEGSECKTFQKTRLSHTRHHTGE 409
Qy 355 KPYHCSI CGARNRPNANKTHSRHISGKPYKCTCGSRFVQRVQSPSPSGFGKPA----- 410
```


Oy 468 SHVRIHTGEKPYHCDPCGLHFRHKSQLRLH 497
Db :||||| | ||| :||
580 MEQRVHTGEKPYTCGACGKHFSQASSLQLH 609

Search completed: August 9, 2005, 14:06:03
Job time : 45 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 13:56:02 ; Search time 177 Seconds
(without alignments)
1498.627 Million cell updates/sec

Title: US-09-815-379-4

Perfect score: 2789

Sequence: 1 MGSPAAPGALGYVRETRH.....ROKHGAATNTKVHILGGP 518

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2556	91.6	480	1	BL6B HUMAN
2	2272	81.5	474	1	BL6B MOUSE
3	1138	40.8	707	1	BC16 MOUSE
4	1138	40.8	707	2	Q8CB25
5	1128.5	40.5	706	1	BC16 HUMAN
6	1125	40.3	701	2	Q6DCT1
7	1084.5	38.9	704	2	Q6PC08
8	575.5	20.6	803	1	ZB17 HUMAN
9	568.5	20.4	756	2	Q8BX2
10	558	20.0	794	1	ZB17 MOUSE
11	542.5	19.5	792	2	Q6GN31
12	539.5	19.3	765	2	Q6ZSB9
13	509.5	18.3	539	1	ZF67 HUMAN
14	505.5	18.1	573	2	Q6BD2
15	491.5	17.6	544	1	ZF67 MOUSE
16	490.5	17.6	706	2	Q8BGY0
17	486.5	17.4	710	2	Q8OX44
18	485.5	17.4	645	2	Q6ZSY6
19	485	17.4	610	2	Q6PLG7
20	482.5	17.3	697	1	ZB34 HUMAN
21	482	17.3	581	2	Q8GZ11
22	482	17.3	610	2	Q9NPC7
23	478	17.1	610	2	Q8GZ12
24	478	17.1	610	2	Q99MD8
25	476	17.1	688	1	HK33 HUMAN
26	472.5	16.9	582	2	Q9CXJ8
27	469.5	16.8	711	1	Z336 HUMAN
28	462.5	16.6	810	2	Q7ZVR6
29	460.5	16.5	688	2	Q7TPC4
30	457.5	16.4	641	2	Q9JMG9
31	455.5	16.3	609	2	Q6DDV0
					Q8nl43 homo sapien
					Q82822 mus musculus
					P41183 mus musculus
					Q8cb25 mus musculus
					Q6dct1 xenopus lae
					Q6pc08 brachydanio
					Q13105 homo sapien
					Q8bxx2 mus musculus
					Q60821 mus musculus
					Q6gn31 xenopus lae
					Q6zsb9 homo sapien
					Q15156 homo sapien
					Q68ar2 homo sapien
					Q64321 mus musculus
					Q8bgy0 m. mus muscu
					Q8ox44 mus musculus
					Q6zsy6 homo sapien
					Q6plg7 mus musculus
					Q43167 homo sapien
					Q86z11 homo sapien
					Q9npc7 homo sapien
					Q86z12 homo sapien
					Q99md8 mus musculus
					P10074 homo sapien
					Q9cxj8 mus musculus
					Q9h16 homo sapien
					Q7zvr6 brachydanio
					Q9jmg9 mus musculus
					Q6ddv0 xenopus lae

32	452	16.2	703	2	Q66JA5	Q66ja5 xenopus lae
33	447	16.0	687	1	Z278 HUMAN	Q9hbel h zinc fing
34	442.5	15.9	703	2	Q9R161	Q9r161 mus musculu
35	441	15.8	470	2	Q9PVP8	Q9pvp8 xenopus lae
36	440.5	15.8	569	1	ZBT7 RAT	Q9z48 ratus norv
37	437.5	15.7	565	1	ZBT7 MOUSE	Q88939 mus musculu
38	436.5	15.7	725	2	Q96SZ4	Q96sz4 homo sapien
39	434.5	15.6	614	1	ZF29 MOUSE	Q07230 mus musculu
40	434	15.6	595	2	Q6ZQY9	Q6zqy9 homo sapien
41	434	15.6	613	2	O727L9	Q7z7l9 homo sapien
42	433	15.5	548	2	Q6AYK0	Q6ayk0 ratus norv
43	432.5	15.5	555	2	Q6Z518	Q6z518 mus musculu
44	430.5	15.4	555	2	Q9Z3D1	Q9z3d1 mus musculu
45	430	15.4	464	2	Q6ZNA0	Q6zna0 homo sapien

ALIGNMENTS

RESULT 1
ID BL6B HUMAN STANDARD; PRT; 480 AA.
AC Q8NL43; Q8PCB4;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE B-cell CLL/lymphoma 6 member B protein (Bcl6-associated zinc finger protein) (Zinc finger protein 62).
GN Name=BCL6B; Synonyms=BAZF, ZNF62;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RX PubMed=11855826; DOI=10.1006/bbrc.2002.6481;
RA Sakashita C., Fukuda T., Okabe S., Kobayashi H., Hirose S., Tokuhisa T., Miyasaka N., Miura O., Miki T.;
RT "Cloning and characterization of the human BAZF gene, a homologue of the BCL6 oncogene."
RL Biochem. Biophys. Res. Commun. 291:567-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Lochewlan N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McQuellan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Wozny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting R., Madan A., Young A.C., Shveychenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Acts as a sequence-specific transcriptional repressor in association with BCL6. May function in a narrow stage or be related to some events in the early B cell development.
CC -!- SUBUNIT: Associates with BCL6 through the BTB domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher expression found in heart and placenta.

CC -!- SIMILARITY: Contains 1 BTF/POZ domain.
 CC -!- SIMILARITY: Contains 5 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB076580; BAC00962.1; -;
 DR EMBL; AB076581; BAC00963.1; -;
 DR EMBL; BC059404; AH59404.1; -;
 DR PIR; JC7812; JC7812.
 DR HSP; P08046; 1G2F.
 DR Genew; HGNC:1002; BCL6B.
 DR InterPro; IPR000210; BTF_POZ.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00651; BTF; 1.
 DR Pfam; PF00096; zF-C2H2; 5.
 DR SMART; SM00225; BTF; 1.
 DR SMART; SM00355; Znf_C2H2; 5.
 DR PROSITE; PS50097; BTF; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
 KW Metal-binding; Nuclear_protein; Repeat; Zinc; Zinc-finger.
 FT DOMAIN 38 105 BTF.
 FT ZN_FING 329 351 C2H2-type 1.
 FT ZN_FING 357 379 C2H2-type 2.
 FT ZN_FING 385 407 C2H2-type 3.
 FT ZN_FING 413 435 C2H2-type 4.
 FT ZN_FING 441 464 C2H2-type 5.
 FT DOMAIN 144 189 Pro-rich.
 FT DOMAIN 234 245 Ser-rich.
 FT CONFLICT 307 307 C -> S (in Ref. 2).
 SQ SEQUENCE 480 AA; 51618 MW; 36F1F01553C375A CRC64;

Query Match 91.6%; Score 2556; DB 1; Length 480;
 Best Local Similarity 92.7%; Pred. No. 1.2e-117;
 Matches 480; Conservative 0; Mismatches 0; Indels 38; Gaps 1;

QY 1 MGSPAAPGALGYVREFTHSSDVLGNLNLRLGILTDVLLVGGQPLRAHKAVALIACS 60
 DB 1 MGSPAAPGALGYVREFTHSSDVLGNLNLRLGILTDVLLVGGQPLRAHKAVALIACS 60
 QY 61 GPFYIFRGRAGVGVVLSLPGGPARGFAPLLDFMTYSLRLSPATAPAVLAATYQLM 120
 DB 61 GPFYIFRGRAGVGVVLSLPGGPARGFAPLLDFMTYSLRLSPATAPAVLAATYQLM 120
 QY 121 EHVQACHRFIOASVEPLGISIRPLEABPTPTAPPGRSPRRSGHPDPPTESRSCSQG 180
 DB 121 EHVQACHRFIOASVEPLGISIRPLEABPTPTAPPGRSPRRSGHPDPPTESRSCSQG 180
 QY 181 PPSASPDPKCNWKKYKIVLNSQASQAGSLVGRSSQPCQARLPSGDEASSSSSS 240
 DB 181 PPSASPDPKCNWKKYKIVLNSQASQAGSLVGRSSQPCQARLPSGDEASSSSSS 240
 QY 241 SSSSEEGIPGPQSRSLPTAATVQKCAPASTPYLLTSQAQDTSGSPSERARLPQSE 300
 DB 241 SSSSEEGIPGPQSRSLPTAATVQKCAPASTPYLLTSQAQDTSGSPSERARLPQSE 300
 QY 301 PFSQCNCAVAGCSGLDVLPGDDEKPKYKCOLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 DB 301 PFSQCNCAVAGCSGLDVLPGDDEKPKYKCOLCRSSFRYKGNLASHRTVHTGKPYHCS 360
 QY 361 ICGARFNPANLKTSHRSHSGKPKYKCTCGSRFVQVRSQPPSGRQKGPARGVGQKGGF 420
 DB 361 ICGARFNPANLKTSHRSHSGKPKYKCTCGSRFVQVRSQPPSGRQKGPARGVGQKGGF 420
 QY 421 CSSQRDLKSPPSQVAHRAHVLHTGKPYPCPTCGTRFRLQTLKSHVRIHTGKPYH 480
 DB 421 CSSQRDLKSPPSQVAHRAHVLHTGKPYPCPTCGTRFRLQTLKSHVRIHTGKPYH 480
 QY 442 QVAHLRAHVLHTGKPYPCPTCGTRFRLQTLKSHVRIHTGKPYH 442

QY 481 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHLGGP 518
 DB 443 CDPGLHFRHKSQRLHLRQKHGAATNTKVHYHLGGP 480

RESULT 2
 BL6B MOUSE
 ID BL6B MOUSE STANDARD; PRT; 474 AA.
 AC 089282; 08CCJ6;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE B-cell CLL/lymphoma 6 member B protein (Bcl6-associated zinc finger protein).
 DE protein.
 GN Name=Bcl6b; Synonyms=BAZF;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, INDUCTION, AND
 RP INTERACTION WITH BCL6.
 RX MEDLINE=98298267; PubMed=9632807;
 RA Okabe S., Fukuda T., Ishibashi K., Kojima S., Okada S., Hatanho M.,
 RA Ebara M., Saisho H., Tokuhisa T.;
 RT "BAZF, a novel Bcl6 homolog, functions as a transcriptional
 RT repressor".
 RL Mol. Cell. Biol. 18:4235-4244 (1998).
 RN [2]
 RP SEQUENCE OF 304-474 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojbori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guenichon S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs".
 RL Nature 420:563-573 (2002).
 RN [3]
 RP FUNCTION.
 RX PubMed=15314041; DOI=10.1093/intimm/dxh144;
 RA Takamori M., Hatanho M., Arima M., Sakamoto A., Fujimura L.,
 RA Hartatik T., Kuriyama T., Tokuhisa T.;
 RT "BAZF is required for activation of naive CD4 T cells by TCR
 RT triggering".
 RL Int. Immunol. 16:1439-1449 (2004).
 CC -!- FUNCTION: Acts as a sequence-specific transcriptional repressor in
 CC association with BCL6. Necessary for activation of naive T cells
 CC to antigenic stimulation. May attenuate the regulatory effect of

CC BCL6 on antigenic activation of naive CD4 T cells by forming an
 CC heterodimer with BCL6.
 CC - SUBUNIT: Associates with BCL6 through the BTB domain.
 CC - SUBCELLULAR LOCATION: Nuclear.
 CC - TISSUE SPECIFICITY: Ubiquitously expressed with higher expression
 CC found in heart and lung.
 CC - INDUCTION: Induced in activated lymphocytes.
 CC - DOMAIN: Amino acids 178-210 are essential for repression activity.
 CC - SIMILARITY: Contains 1 BTB/POZ domain.
 CC - SIMILARITY: Contains 5 C2H2-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB011665; BAA31223.1; -
 CC EMBL; AK032647; BAC27970.1; -
 CC HSPD; P08046; IGZF.
 CC MGD; MGI:1278332; Bcl6b.
 CC GO; GO:0005634; C:nucleus; IDA.
 CC GO; GO:0003677; F:DNA binding; IDA.
 CC GO; GO:0005515; F:protein binding; IPI.
 CC GO; GO:0016564; F:transcriptional repressor activity; IDA.
 CC GO; GO:0000122; P:negative regulation of transcription from P. . .; IDA.
 CC InterPro; IPR000210; BTB_POZ.
 CC InterPro; IPR007087; Znf_C2H2.
 CC Pfam; PF00651; BTB; 1.
 CC Pfam; PF00096; zf-C2H2; 5.
 CC SMART; SM00225; BTB; 1.
 CC SMART; SM00355; Znf_C2H2; 5.
 CC PROSITE; PS00097; BTB; 1.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
 CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
 CC Metal-binding; Nuclear protein; Repeat; Zinc; Zinc-finger.
 KW DOMAIN 38 105
 FT ZN_FING 323 345
 FT ZN_FING 351 373
 FT ZN_FING 379 401
 FT ZN_FING 407 429
 FT ZN_FING 435 458
 FT ZN_FING 144 189
 FT DOMAIN 234 239
 FT DOMAIN 474 AA; 51339 MW; 94886AB265B018D8 CRC64;
 SQ SEQUENCE 474 AA; 51339 MW; 94886AB265B018D8 CRC64;
 Query Match 81.5%; Score 2272; DB 1; Length 474;
 Best Local Similarity 83.4%; Pred.No. 9.1e-104;
 Matches 432; Conservative 11; Mismatches 31; Indels 44; Gaps 3;
 QY 1 MGSPAAPGALGVYREFTRHSDVNLGNELRLGILTDVLLVGGQPLRAHKAFLIACS 60
 DB 1 MGSPAAPGALGVYREFTRHSDVNLGNELRLGILTDVLLVGGQPLRAHKAFLIACS 60
 QY 61 GFFYSIFRAGAGVDVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPVLAATYIQM 120
 DB 61 GFFYSIFRAGAGVDVLSLPGGPEARGFAPLLDFMYTSRLRLSPATAPVLAATYIQM 120
 QY 121 EHVQACHRFQASVEPIGILSRPLEAEPPTPTAPPGSPRRRGGHDPPTESRSCSQ 180
 DB 121 EHVQACHRFQASVEPIGILSRPLEAEPPTPTAPPGSPRRRGGHDPPTESRSCSQ 180
 QY 181 PPSPASPPKACNKKYKIVLNSQASGLVGRSSGQPCQARLPSPGDEASSSSSS 240
 DB 181 SPSPASPPKACNKKYKIVLNSQASGLVGRSSGQPCQARLPSPGDEA-----CS 234
 QY 241 SSSSSEEGFIPGOSRLSPTAATVQKCAPASTPYLLTSQAQDTSPSPERAPLPGE 300
 DB 235 SSSSSEEGTFGLQSRSLATTTAFKCGALANNYSYLPTRAQETSLPASKQANPPGSE 294
 QY 301 FFSQNCBAVAGCSGLDVLPGDEDKPKYKQLCRSAPRYKGNLASHRTVHTGKPYCS 360

DB 295 FFSQNCBAVAGCSGLDVLPGDEDKPKYKQLCRSAPRYKGNLASHRTVHTGKPYCS 354
 QY 361 ICGARFNRPANLKHSHSGEKPKYCETCGSRFQVRSQPPSGFQKPGAGVGQKGF 420
 DB 355 ICGARFNRPANLKHSHSGEKPKYCETCGSRF----- 389
 QY 421 CSSQRODLKPPSQVAHLRAHVLHITGKEPKYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 480
 DB 390 -----QVAHLRAHVLHITGKEPKYPCPTCGTRFRHLQTLKSHVRIHTGKPYH 436
 QY 481 CDPCGLHFRHKSQRLHLRQKHGAATNTKVHYHILGCP 518
 DB 437 CDPCGLHFRHKSQRLHLRQKHGAATNTKVHYHILGCP 474
 RESULT 3
 BCL6 MOUSE
 ID BCL6 MOUSE STANDARD; PRT; 707 AA.
 AC P41183; O61065;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE B-cell lymphoma 6 protein homolog.
 GN Name=Bcl6; Synonyms=Bcl-6;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Skeletal muscle;
 RX MEDLINE=96038894; PubMed=7478591;
 RA Fukuda T., Miki T., Yoshida T., Hatano M., Ohashi K., Hirose S.,
 RA Tokuhisa T.;
 RT "The murine BCL6 gene is induced in activated lymphocytes as an
 RT immediate early gene.";
 RL Oncogene 11:1657-1663(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=96247530; PubMed=8652841;
 RA Allman D., Jain A., Dent A., Maile R.R., Selvaggi T., Kehry M.R.,
 RA Staudt L.M.;
 RT "BCL-6 expression during B-cell activation.";
 RL Blood 87:5257-5268(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Olfactory epithelium;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP INTERACTION WITH BCL6B.
 RX MEDLINE=98298267; PubMed=9632807;

RA Okabe S., Fukuda T., Ishibashi K., Kojima S., Okada S., Hatano M.,
 RA Ebata M., Saisho H., Tokuhisa T.,
 RT "BAZF, a novel Bcl6 homolog, functions as a transcriptional
 RT repressor."; Mol. Cell. Biol. 18:4235-4244(1998).
 RL [5]
 RN INTERACTION WITH ZBTB7, AND SUBCELLULAR LOCATION.
 RX MEDLINE=99124383; PubMed=9927193; DOI=10.1038/sj.onc.1202332;
 RA Davies J.M., Howe N., Kabatowski J., Huang Q.-H., Zhu J., Brand N.J.,
 RA Leprince D., Dhordain P., Cook M., Moriss-Kay G., Zelent A.;
 RT "Novel BTF/POZ domain zinc-finger protein, LRF, is a potential target
 RT of the LAZ-3/BCL-6 oncogene."; Oncogene 18:365-375(1999).
 RL [5]
 RN FUNCTION: Transcription regulator that probably play an important
 CC role in lymphomagenesis.
 CC - SUBUNIT: Interacts with ZBTB7 and BCL6.
 CC - SUBCELLULAR LOCATION: Nuclear.
 CC - SIMILARITY: Contains 1 BTF/POZ domain.
 CC - SIMILARITY: Contains 6 C2H2-type zinc fingers.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D38377; BAA07456.1; -
 DR EMBL; U41465; AAB17432.1; -
 DR EMBL; BC052315; AAB52315.1; -
 DR HSP; P08046; IG2F.
 DR MGD; MGI:107187; Bcl6
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0016564; F:transcriptional repressor activity; IDA.
 DR GO; GO:0000122; P:negative regulation of transcription from P. .; IDA.
 DR InterPro; IPR000210; BTF POZ.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00651; BTF; 1.
 DR Pfam; PF00096; zf-C2H2; 6.
 DR PROSITE; PS00097; BTF; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 6.
 DR Activator; DNA-binding, Metal-binding, Nuclear protein; Repeat;
 KW Transcription regulation; Zinc-finger.
 FT DOMAIN 32 99
 FT ZN_FING 519 542 C2H2-type 1.
 FT ZN_FING 547 569 C2H2-type 2.
 FT ZN_FING 575 597 C2H2-type 3.
 FT ZN_FING 603 625 C2H2-type 4.
 FT ZN_FING 631 653 C2H2-type 5.
 FT ZN_FING 659 682 C2H2-type 6.
 FT CONFLICT 456 456 A -> G (in Ref. 2).
 SQ SEQUENCE 707 AA; 78981 MW; 2051DD808D32D5EC CRC64;
 Query Match 40.8%; Score 1138; DB 1; Length 707;
 Best Local Similarity 36.4%; Pred. No. 3.9e-48;
 Matches 269; Conservative 51; Mismatches 149; Indels 270; Gaps 14;
 QY 1 MGSPAAGALGVYREFTRHSDVNLNLRILGILTDVLLVGGQPLRAHKAIVLIACS 60
 DB 1 MASPA--DSCI---QFTHASDVLLNLRSLRDLTDVVIVVSREQFRAHKTIVLACS 54
 QY 61 GPFFSIFRGRAGVGVVLSLPGGPGRAGFAPLLDFMYTSRLRLSPATPAVLAATYLOM 120
 DB 55 GLFYSIFTDLKCNLSVINLDEISPEGFCILLDFMYTSRLNLRGINMAVMTTAMVLOM 114
 QY 121 EHVVOACHRFIOASVEPLGISLRPLEAE----- 148
 DB 115 EHVVDTCRKFIKASAEAPALKPPREBFNLNLRMLPHDIMAAYRGREVVNMPLRNTPG 174
 QY 149 -----PPT----- 151

Db 175 CESRAFAPPLYSLGSLTTPPASYPMYSHLPLSTFLFSDLELRDAPRMPVAVPFPKERALPCD 234
 QY 152 ----- 151
 Db 235 SARQVNPVSRPAMEVSPSLCHSNIVSPKEAVPEEARSDIHSVPEGPKAVPSPARNAPY 294
 QY 152 -----PPTAP-----PPGSPRRSEGHDPDPTTSRS----- 176
 Db 295 FPCDKAKSEERPSSEDEIALHFEPNAPLNRRKGLVSPQSPQSKDCQPNPTSCSSKNA 354
 QY 177 -----CSGGPPSPAPDPKACNWKYKIVILNS-----QASQAGSLVGRSSQCPQQA 225
 Db 355 CILQASGSPPAKFTDFKACNWKYKIVILNSQNAKPEGSEQAELGRLSPRAYPAPPA 414
 QY 226 RLPSGDEAS--SSSSSSSSSEEGPIPGQPSRL-----SPTAATVQF----- 266
 Db 415 CQPMPEANLDLOSPTKLSAGSDSTIP-QASRLNLDVNRSLAGSPRSSSESHPLYMHP 473
 QY 267 -KCGAPASTPYLLTSAQDTSQSPSERARPLPQSEF-----FSCNCEAVAGCSSG 316
 Db 474 PKTSCGSGSQSPQHTMCLHTAGTFFPEEMGETQSEYSDSSCENGTFPCNECDRFSSEAS 533
 QY 317 LD-SLVPGDDEKPKCOLCRSSFRYKGNLASHETVHTGKPYHCSICGABFNRPANLKT 375
 Db 534 LKRHTLQTHSDKPKCDRCQASFRYKGNLASHETVHTGKPYHCSICGABFNRPANLKT 593
 QY 376 SRTHSGKPKYKCTCGSRFVQVRSQPPSGFGQKPARGVGQKGFCSQQDLKSPSPQV 435
 Db 594 TRIHSGKPKYKCTCGARFV-----QV 615
 QY 436 AHLRAHVLIHTGKPYKPCPTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKQLR 495
 Db 616 AHLRAHVLIHTGKPYKPCPTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKQLR 675
 QY 496 LHLRQKHGAATNTKVYHI 514
 Db 676 LHLRQKHGAITNTKVQYRV 694
 RESULT 4
 Q8CB25 PRELIMINARY; PRT; 707 AA.
 AC Q8CB25; 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Mus musculus adult female vagina cDNA, RIKEN full-length enriched
 DE library, clone:9930032A10 product:B-cell leukemia/lymphoma 6, full
 DE insert sequence.
 GN Name=Bcl6;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Vagina;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Vagina;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Vagina;
 RA The FANTOM Consortium,

RA Ye B.H., Lieta F., Lo Coco F., Knowles D.M., Offit K.,
RA Chaganti R.S.K., Dalla-Favera R.;
RT "Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-
RT cell lymphoma";
RL Science 262:747-750(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94100541; PubMed=8274740;
RA Miki T., Kawanata N., Hirosewa S., Aoki N.;
RT "Gene involved in the 3q27 translocation associated with B-cell
RT lymphoma, BCL6, encodes a Kruppel-like zinc-finger protein.";
RL Blood 83:26-32(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281738; PubMed=8506375;
RA Baron B.W., Nucifora G., McCabe N., Espinosa R. III, le Beau M.M.,
RA McKeithan T.W.;
RT "Identification of the gene associated with the recurring chromosomal
RT translocations t(3;14)(q27;q32) and t(3;22)(q27;q11) in B-cell
RT lymphomas";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5262-5266(1993).
CC -!- FUNCTION: Transcriptional regulator that probably plays an
CC important role in lymphomagenesis.
CC -!- SUBUNIT: Interacts with ZBTB7 and BCL6 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- DISEASE: Involved in a form of B-cell non-Hodgkin's lymphoma
CC characterized by chromosomal translocation t(3;14)(q27;q32) and
CC t(3;22)(q27;q11) that involves BCL6 and immunoglobulin gene
CC regions.
CC -!- DISEASE: Involved in a form of B-cell leukemia characterized by a
CC chromosomal translocation t(3;11)(q27;q23) that involves BCL6 and
CC POU2AF1 (OBF1).
CC -!- DISEASE: Involved in a t(3;4)(q27;p11) chromosomal translocation
CC with ARHG (TTF).
CC -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC -!- SIMILARITY: Contains 6 C2H2-type zinc fingers.
CC -!- DATABASE: NAME=Atlas Genet. Cytoenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/BCL6ID20.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J21943; CAA79937.1; -;
DR EMBL; U00115; AAC50054.1; -;
DR EMBL; S67779; -; NOT_ANNOTATED_CDS.
DR FIR; I52586; I52586.
DR PDB; 1R28; X-ray; A/B=3-129.
DR PDB; 1R29; X-ray; A=3-129.
DR PDB; 1R2B; X-ray; A/B=3-129.
DR TRANSFAC; T02322; -;
DR Genew; HGNC:1001; BCL6.
DR MIM; 109565; -;
DR GO; GO:0008954; P:inflammatory response; TAS.
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00651; BTB; 1.
DR Pfam; PF00096; zf-C2H2; 6.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00355; Znf_C2H2; 6.
DR PROSITE; PS50097; BTB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
KW 3D-structure; Activator; Chromosomal translocation; DNA-binding;
KW Metal-binding; Nuclear protein; Polymorphism; Proto-oncogene; Repeat;
KW Transcription regulation; Zinc-finger.

FT	DOMAIN	32	99	BTB.
FT	ZN_FING	518	541	C2H2-type 1.
FT	ZN_FING	541	568	C2H2-type 2.
FT	ZN_FING	574	596	C2H2-type 3.
FT	ZN_FING	602	624	C2H2-type 4.
FT	ZN_FING	630	652	C2H2-type 5.
FT	ZN_FING	658	681	C2H2-type 6.
FT	VARIANT	676	676	H -> Y (in dbSNP:1056936).
FT	CONFLICT	347	347	/FTID=VAR_014825.
FT	CONFLICT	393	393	S -> A (in Ref. 2).
FT	CONFLICT	498	498	E -> G (in Ref. 2).
FT	CONFLICT	498	498	P -> A (in Ref. 3).
SQ	SEQUENCE	706 AA;	78846 MW;	E38D83C213DAE2D0 CRC64;

Query Match 40.5%; Score 1128.5; DB 1; Length 706;
Best Local Similarity 35.0%; Pred. No. 1.1e-47;
Matches 270; Conservative 48; Mismatches 117; Indels 337; Gaps 16;

Qy	1	MGSPAAEGALGVRETRHSSDVLGNLRLRGILTDVLLVGGQPLRAHKAVALIACS	60
Db <td>1<td>MASPA--DSCI---QFTHASDVLNLRSLRSDILTDVIVVRSREOFRAHKTIVLMACS</td><td>54</td></td>	1 <td>MASPA--DSCI---QFTHASDVLNLRSLRSDILTDVIVVRSREOFRAHKTIVLMACS</td> <td>54</td>	MASPA--DSCI---QFTHASDVLNLRSLRSDILTDVIVVRSREOFRAHKTIVLMACS	54
Qy <td>61<td>GFYYSIFRGAGVGVVLSLPGGPEARGFAPLDDFMYTSRLRLSPATAPAVLAAATYLOM</td><td>120</td></td>	61 <td>GFYYSIFRGAGVGVVLSLPGGPEARGFAPLDDFMYTSRLRLSPATAPAVLAAATYLOM</td> <td>120</td>	GFYYSIFRGAGVGVVLSLPGGPEARGFAPLDDFMYTSRLRLSPATAPAVLAAATYLOM	120
Db <td>55<td>GLFVSIITDQKMLSVINLDPEINPEGFCILLDFMYTSRLNLRGNIMAVMATYLOM</td><td>114</td></td>	55 <td>GLFVSIITDQKMLSVINLDPEINPEGFCILLDFMYTSRLNLRGNIMAVMATYLOM</td> <td>114</td>	GLFVSIITDQKMLSVINLDPEINPEGFCILLDFMYTSRLNLRGNIMAVMATYLOM	114
Qy <td>121<td>EHVVQACHRFIOAS-----</td><td>134</td></td>	121 <td>EHVVQACHRFIOAS-----</td> <td>134</td>	EHVVQACHRFIOAS-----	134
Db <td>115<td>EHVVDTCKRIKASEAEWMSAIKPPREFFLNSRLMPODIMAYRGREVVNNPLRSAPG</td><td>174</td></td>	115 <td>EHVVDTCKRIKASEAEWMSAIKPPREFFLNSRLMPODIMAYRGREVVNNPLRSAPG</td> <td>174</td>	EHVVDTCKRIKASEAEWMSAIKPPREFFLNSRLMPODIMAYRGREVVNNPLRSAPG	174
Qy <td>135</td> <td>-----YELGIS-----</td> <td>141</td>	135	-----YELGIS-----	141
Db <td>175<td>CESRAFAFSLYGLSTPPASYSMYSLPVSLLFSDEFRDVRMPVANPPKERALPCDS</td><td>234</td></td>	175 <td>CESRAFAFSLYGLSTPPASYSMYSLPVSLLFSDEFRDVRMPVANPPKERALPCDS</td> <td>234</td>	CESRAFAFSLYGLSTPPASYSMYSLPVSLLFSDEFRDVRMPVANPPKERALPCDS	234
Qy <td>142</td> <td>-----LRPL-----</td> <td>145</td>	142	-----LRPL-----	145
Db <td>235<td>ARVPGEYSRPTLEVSFNVCHSNIYSPKTIPEARSDMHYSVAEGLKPAAPARNAPYF</td><td>294</td></td>	235 <td>ARVPGEYSRPTLEVSFNVCHSNIYSPKTIPEARSDMHYSVAEGLKPAAPARNAPYF</td> <td>294</td>	ARVPGEYSRPTLEVSFNVCHSNIYSPKTIPEARSDMHYSVAEGLKPAAPARNAPYF	294
Qy <td>146</td> <td>-----EABPPT-----PPTAP-----PPGSPRRSEGHDPPTTES-</td> <td>176</td>	146	-----EABPPT-----PPTAP-----PPGSPRRSEGHDPPTTES-	176
Db <td>295<td>PCDKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSKSCQPNSTPSCSKNAC</td><td>354</td></td>	295 <td>PCDKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSKSCQPNSTPSCSKNAC</td> <td>354</td>	PCDKASKEERPSSEDEIALHFEPPNAPLNKGLVSPQSKSCQPNSTPSCSKNAC	354
Qy <td>177</td> <td>---CSQPPPPSPADPKACWKKYKIVLNSQASQSLGERSGQPCPOARLPBGDEA</td> <td>233</td>	177	---CSQPPPPSPADPKACWKKYKIVLNSQASQSLGERSGQPCPOARLPBGDEA	233
Db <td>355<td>ILQASGSPPAKSPTPDKACWKKYKIVLN-----</td><td>384</td></td>	355 <td>ILQASGSPPAKSPTPDKACWKKYKIVLN-----</td> <td>384</td>	ILQASGSPPAKSPTPDKACWKKYKIVLN-----	384
Qy <td>234<td>SSSSSSSSSSSEEGPIPGQSRSLSPATAATVQFKGAP-----ASTPYLLTQAQDT-</td><td>285</td></td>	234 <td>SSSSSSSSSSSEEGPIPGQSRSLSPATAATVQFKGAP-----ASTPYLLTQAQDT-</td> <td>285</td>	SSSSSSSSSSSEEGPIPGQSRSLSPATAATVQFKGAP-----ASTPYLLTQAQDT-	285
Db <td>385</td> <td>-----SUNQNAKPEGPEQAEGLRUSPRAYTAPACQPPMEPNLDLOSPTKLSAGEDST</td> <td>439</td>	385	-----SUNQNAKPEGPEQAEGLRUSPRAYTAPACQPPMEPNLDLOSPTKLSAGEDST	439
Qy <td>286</td> <td>-----SGSP-----SERARPL-----</td> <td>296</td>	286	-----SGSP-----SERARPL-----	296
Db <td>440<td>IPQASRLNINVRSMTCSPRSSSESHPLMYHPKCTSCGSQSPQHAEMCLHTAGTPPE</td><td>499</td></td>	440 <td>IPQASRLNINVRSMTCSPRSSSESHPLMYHPKCTSCGSQSPQHAEMCLHTAGTPPE</td> <td>499</td>	IPQASRLNINVRSMTCSPRSSSESHPLMYHPKCTSCGSQSPQHAEMCLHTAGTPPE	499
Qy <td>297</td> <td>-----PGSEF-----FSCQCEAVAGSSGLD-SLVPGDEKPKYKQLCRSEFYKG</td> <td>342</td>	297	-----PGSEF-----FSCQCEAVAGSSGLD-SLVPGDEKPKYKQLCRSEFYKG	342
Db <td>500<td>EMGETQSEYSDSSCENGAFPCNECDRCFSEASLKRHRTLQTHSKDKYKCDRCQASFRYKG</td><td>559</td></td>	500 <td>EMGETQSEYSDSSCENGAFPCNECDRCFSEASLKRHRTLQTHSKDKYKCDRCQASFRYKG</td> <td>559</td>	EMGETQSEYSDSSCENGAFPCNECDRCFSEASLKRHRTLQTHSKDKYKCDRCQASFRYKG	559
Qy <td>343<td>NLASHRTVHTGKPYHCSIICARNRPNALKTHSRHSKPYKCYKTCGSRFVQVRSQPP</td><td>402</td></td>	343 <td>NLASHRTVHTGKPYHCSIICARNRPNALKTHSRHSKPYKCYKTCGSRFVQVRSQPP</td> <td>402</td>	NLASHRTVHTGKPYHCSIICARNRPNALKTHSRHSKPYKCYKTCGSRFVQVRSQPP	402
Db <td>560<td>NLASHKTVHTGKPYRCNICQANRPNALKTHSRHSKPYKCYKTCGARFV-----</td><td>612</td></td>	560 <td>NLASHKTVHTGKPYRCNICQANRPNALKTHSRHSKPYKCYKTCGARFV-----</td> <td>612</td>	NLASHKTVHTGKPYRCNICQANRPNALKTHSRHSKPYKCYKTCGARFV-----	612
Qy <td>403<td>SGFOGKPARGVGKGKGGFCSSQRQDLKSPPSQVAHLRAHVLHTGKPYPCPTCTPRFH</td><td>462</td></td>	403 <td>SGFOGKPARGVGKGKGGFCSSQRQDLKSPPSQVAHLRAHVLHTGKPYPCPTCTPRFH</td> <td>462</td>	SGFOGKPARGVGKGKGGFCSSQRQDLKSPPSQVAHLRAHVLHTGKPYPCPTCTPRFH	462
Db <td>613</td> <td>-----QVAHLRAHVLHTGKPYPCPTCTPRFH</td> <td>641</td>	613	-----QVAHLRAHVLHTGKPYPCPTCTPRFH	641
Qy <td>463<td>LQTLKSHVRIHTGKPYHCDPCGLHFRHKQLRLHLRQKGAATNTKVHYH</td><td>514</td></td>	463 <td>LQTLKSHVRIHTGKPYHCDPCGLHFRHKQLRLHLRQKGAATNTKVHYH</td> <td>514</td>	LQTLKSHVRIHTGKPYHCDPCGLHFRHKQLRLHLRQKGAATNTKVHYH	514
Db <td>642<td>LQTLKSHLRIHTGKPYHCEKCNLHFRHKQLRLHLRQKGAITNTKVQYRV</td><td>693</td></td>	642 <td>LQTLKSHLRIHTGKPYHCEKCNLHFRHKQLRLHLRQKGAITNTKVQYRV</td> <td>693</td>	LQTLKSHLRIHTGKPYHCEKCNLHFRHKQLRLHLRQKGAITNTKVQYRV	693

RESULT 6
Q6DCT1

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U20647; AAC50256.1; --
 EMBL; AL034555; CAB85445.1; --
 EMBL; Y03723; CAA70889.1; --
 EMBL; M88369; AAG61327.1; --
 PIR; D45193; D45193.
 PIR; I38940; I38940.
 HSP; P07248; IARE.
 TRANSFAC; T03414; --
 Genew; HGNC:12936; ZBTB17.
 MIM; 604084; --
 InterPro; IPR000210; BTB_POZ.
 InterPro; IPR007087; Znf_C2H2.
 Pfam; PF00651; BTB; 1.
 Pfam; PF00096; zf-C2H2; 13.
 ProDom; PD000003; Znf_C2H2; 3.
 SMART; SM00225; BTB; 1.
 SMART; SM00355; Znf_C2H2; 13.
 PROSITE; PS00097; BTB; 1.
 PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
 PROSITE; PS00157; ZINC_FINGER_C2H2_2; 13.
 DNA-binding; Metal-binding; Nuclear protein; Repeat;
 Transcription regulation; Zinc-finger.
 DOMAIN 1 104 BTB.
 FT ZN_FING 306 328 C2H2-type 1.
 FT ZN_FING 334 356 C2H2-type 2.
 FT ZN_FING 362 384 C2H2-type 3.
 FT ZN_FING 390 412 C2H2-type 4.
 FT ZN_FING 418 440 C2H2-type 5.
 FT ZN_FING 446 468 C2H2-type 6.
 FT ZN_FING 474 496 C2H2-type 7.
 FT ZN_FING 502 524 C2H2-type 8.
 FT ZN_FING 530 552 C2H2-type 9.
 FT ZN_FING 558 580 C2H2-type 10.
 FT ZN_FING 586 608 C2H2-type 11.
 FT ZN_FING 614 637 C2H2-type 12.
 FT ZN_FING 717 739 C2H2-type 13.
 FT CONFLICT 73 73 M -> V (in Ref. 2).
 SEQUENCE 803 AA; 87959 MW; FEE72BA4E14AE200 CRC64;

Query Match 20.6%; Score 575.5; DB 1; Length 803;
 Best Local Similarity 29.4%; Pred. No. 1.5e-20;
 Matches 180; Conservative 64; Mismatches 205; Indels 163; Gaps 20;
 16 EFTRHSSDVILNLELRGLITDVLVGGQPLRAKAVLIACSGFFVIFRGRAG-VG 74
 2 DFPQHSQHVLEQLNQORQLGLLDCDTFFVGVGHFAKRAVLAACSEYFKMLFVDQVVH 61
 75 VDVLSPGPPARGAPALIDFMYTSRLSLPATAPAVLAATAATYLMQEHVVOACHRFIOAS 134
 62 LDI ----SNRAGLGQMLEFMYTAKLSLSPENVDDVLAVATFLQMDIITACH----- 109
 135 YEPLGISLRLEAEPTTP-----TAPPGSPRSEGHPPDPPTESRCSQGPSPASPDP 188
 110 -----ALKSL-AEPATSPGGNAEALATGGDKRAKEKVATSTLSRLEQAGRSTPIGP- 161
 199 PKACNKKYKIVLNQASQAGSLVGERSSGQPCQARLP-----SGDEASSSSSSSSSS 245
 162 ---SRDLKEERGQQAASAGAEQTEKADAPREPPVPELVKPDPTSGMAAAEAFALESSE 219
 246 EGPPI-PGQSRSLSPATAVQFKCA-PAST-----PYLLTQSA-----QDTSGSPS--- 290
 220 QEMEVEPARKBEEQKEQEEGAGPAVKEGSGLENGEAPRENENESSAGTDSQGE 279
 291 --ERARPLP-----GSEF-----FSCQNC 307
 280 LGSEARGLRSYTGRTESKAYGSIHKCEDCGKEFTHTGNFKRIHTGKPFSCREC 339

308 EAV-----AGCGSLDLSLVPGDEDKPKYKCOLCRSFRY----- 340
 340 SKAFSDPAACKAEKTHSP---LKPYGCEBGCKSYRLISLNLHLKGRHSGEARYCEDCG 396
 341 -----KGNLASHRTVHTGKPKYHCSICG-----ARFN 367
 397 KLFTTGNLKRHLQVHSGEPYQCYDGRSFSPTSKWRHLETHDTHDKHKCPHCDKKFN 456
 368 RPAHLKTHSIHSGEKPCKCTCGSRFVQVRSPQPGFQGGKPGVGVQKGGFSSQRQD 427
 457 QVGNLKAHLKIHIADGLPKCREGCKQFT-----TSGNLKRHLRIHSGEKPVCVICHQCRQ 510
 428 LKSPPSQVAHLRAHVLHTGKPKYPCPTGTRFRLQTLKSHVRIHTGKPKYHCDPCGLH 487
 511 FADPGA-----LQRHVRHTGKPKQCYVMCGKAFQTQASSLIAHVROHTGKPKYVCERCGR 566
 488 FRHKSQRLHLR 499
 567 FVQSSQLANHIR 578
 RESULT 9
 Q8BX2 PRELIMINARY; PRT; 756 AA.
 ID Q8BX2
 AC Q8BX2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
 DE enriched library, clone:A73050C18 product:weakly similar to ZINC
 DE FINGER PROTEIN 177.
 GN Name=Zfp509;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RA "High-efficiency full-length cDNA cloning";
 RL Meth. Enzymol. 303:19-44(1999).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RL [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RL [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes";
 RL Genome Res. 10:1617-1630(2000).
 RL [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20530913; PubMed=11078861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara B., Watahiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tonaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK043044; BAC31447.1; -;
DR HSSP; P08047; 1SP2.
DR MGD; MGI:1922329; Zfp509.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00651; BTB; 1.
DR Pfam; PF00096; zf-C2H2; 7.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00355; Znf_C2H2; 7.
DR PROSITE; PS00097; BTB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 7.
SQ SEQUENCE 756 AA; 83090 MW; 0D7E6261C61F3225 CRC64;

Query Match 20.4%; Score 568.5; DB 2; Length 756;
Best Local Similarity 27.5%; Pred. No. 3e-20;
Matches 168; Conservative 70; Mismatches 210; Indels 163; Gaps 18;

QY 20 HSDVLGNLNLRLGIITDVTLLVGQPLRAHKAVALIACGFFYIFRGRAGVGVDLVS 79
DB 7 HSCILLQQLREQRIGLLDCMLVVRGVCFAHKNVLAAFSQYFRSLFPQNSSSQKNDVPH 66

QY 80 LPGAPEARGFAPLLDFMTYSLRLSPATAPAVLAATVLMHVQACHRFIOASYEPLG 139
DB 67 L-DVTNVSIGIQLDFMTYSLRLDNLQNIQWMLDTAQLQVQNLNLCHTFLKA----- 120

QY 140 ISLRPLEAEPPTPTAPPFGSP-----RRSEGHDPPTESRSCSQGPPS 183
DB 121 -----PAQLPGIPPCAGGFSLSQVALDGTCAVSEHYF-PFSLQECPEVGQA 166

QY 184 --PASDPKA--CNWKYKIVLNSQASQASGLVGRSSGP----- 221
DB 167 KVAEVRNARAPSANPSRTEV---SKPDAAGSCPELPCKOPNHYKILRTLYSKYQKQT 223

QY 222 -CPQARLPSGEAESSSSSSSSSEEGPIG-----POSRLSPTAATVQFKCG 269
DB 224 ACPQSVQATQPLTKRSASTDLAAADSPPVGEPRPAVLETPHELPSTFVAPVRNNGNSE 283

QY 270 A-PASTP-----YLLTSQAQDTSGSP-----SERARPLFGSEFF 302
DB 284 ADPLSEPAKQMLKKAHMLKKNLFLKSKQQAQSAECTSHPEPDNGLARREESAATEDAVERA 343

QY 303 SCQNC-----AVAGSSGLDSLVPG-----DEDKPKQLCRSSFRYKGNLAS 346
DB 344 GSQTAEKGRGELGPSSREBELPGAPASWEDPDSQALQPKYACELCKPFPKPSNLEL 403

QY 347 HRTVHTGKPYHCICGARFNRPANLKTHSRHSGEKPYKCTCGSRFVQVRSOPPSGQ 406
DB 404 HKRSHTGKPEPCNICGKHFSQAGNLQTHLRHSGEKPYICEICGRFA----- 452

QY 407 GKPARGGV-----GQKGFSCSQDPLKSPSQVAHLRAHV----- 442
DB 453 ---ASGDVQRHIIHSGEKPHLCDTCGRGFSNFGNLKHKHTHTADKVFTCDCECKSPNM 509

QY 443 -----LIHTGKPYPCPTCGFRFRHLQTLKSHVRIHTGKPKVHCDPCGLHFRHKSQ 494
DB 510 QRLVKRVHRTGTERPISCPACGKCFGSGDLRRHVRHTHTGKPYSCVCKCFTRSAVL 569

QY 495 RLHLRQKHGA 505
DB 570 RRHKRM-HGRA 579

RESULT 10
ZB17 MOUSE
ID ZB17_MOUSE STANDARD; PRT; 794 AA.
AC Q60821; Q60699;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE zinc finger and BTB domain containing protein 17 (Zinc finger protein 151) (Zinc finger protein 100) (Zfp-100) (Polyomavirus late initiator promoter binding protein) (LP-1) (Zinc finger protein Z13).
GN Name=Zbtb17; Synonyms=zfp100, Znf151;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Rapp J., Carmichael G.G.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA; TISSUE=Kidney;
RX MEDLINE=96003919; PubMed=7575457;
RA Schulz T.C., Hopwood B., Rathjen P.D., Wells J.R.E.;
RT "An unusual arrangement of 13 zinc fingers in the vertebrate gene Z13.";
RT Z13.";
RL Biochem. J. 311:219-224(1995).
CC -!- FUNCTION: May function as a housekeeping DNA-binding protein that regulates the expression of specific genes.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Found in all the embryonic and adult tissues examined.
CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein family.
CC -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC -!- SIMILARITY: Contains 13 C2H2-type zinc fingers.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/) or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U22396; AAA64848.1; -;
CC EMBL; U14556; AAA85493.1; -;
CC PIR; S59069; S59069.
CC HSSP; P08046; 1GZF.
CC MGD; MGI:107410; Zbtb17.
CC GO; GO:0003677; F:DNA binding; IEA.
CC GO; GO:0007398; P:ectoderm development; IMP.
CC GO; GO:0010003; P:gastrulation (sensu Mammalia); IMP.
CC InterPro; IPR000210; BTB_POZ.
CC InterPro; IPR007087; Znf_C2H2.

Query Match 19.5%; Score 542.5; DB 2; Length 792;
 Best Local Similarity 27.9%; Pred. No. 5.9e-19;
 Matches 178; Conservative 70; Mismatches 228; Indels 163; Gaps 19;

QY 20 HSSDVLGNLRLRGLITDVTLLVGGQPLRAHKAVALIACSGFFYSIFRGRAGVGDVLS 79
 DB 7 HSCILLOLHQRIQGLLDCMLVGVKVRFAKRVNLAAPSOYFRSLFQNSSSQKNDVPH 66
 QY 80 LPGAPEARGFAPLLDFMYTSRLRLSPATAPAVLAATAATYLMQEHVVOACHRFIQAS 139
 DB 67 L-DIKNIGGIGQILDPMYTSRLRLSHLDLNDNVQVMDVAQCLOVQNVLSMCSYFLKLANPVEP 125
 QY 140 ISLRPLEAEPPTAPPSPRRSEGHDPD-----TESR-----SCSQGP-----S 183
 DB 126 ASSTPCSGPLSPFCGVFNPSNQQLSPPLQDIVTEHRHNVVQTLKETAPLNEVS 185
 QY 184 PASDPP-----KACN-----WKYI-----VLNSOAS-----OAGSLVGRS 217
 DB 186 KAALTPLNEMKEIQKHLNQPKLDFYKLFYKEIADKVLQQAAPSLQASLNLEQA 245
 QY 218 SGQPCQA-----RLPSGDEASSSSSSSSSSSS----- 245
 DB 246 AAPTLQAVSTNLQSAVNLQEAASSTNLQEAASSTNLQEAASSTNLQEAASSTNLQEA 305
 QY 246 -----EEGRIPGQSRISP-----TAAT-----VQKCGAPAS 273
 DB 306 AETQPMDFNHSAPLEPSKSLAPLEPLFAFHTFVTNADSLAKPFEQMLKKAHLK 365
 QY 274 TPYLTSQAQDTSGPSRRAPLPQSEFPSCQCEAVAGCSGL-----DS 319
 DB 366 KLMRLSQKAEGYSQPENALQSTESVCSGKVKLEAHDRLLEKESAGIFESIGDT 423
 QY 320 LVPGD-----EDPKYKQLCRSSFRYKGNLASHRTVHTGEKPYHCSICGARPNRPNALKTH 375
 DB 424 LAPAEGVISTNKHVFCDCIGKGRFHRPSNLEQKHSHTGEKPFECSCGKHPQAGNLQTH 483
 QY 376 SRIHSGEKPKECTCGSRFVQVRQPSGFGCKPARGGV-----GQKGFCSQSRQDLKSP 432
 DB 484 LRRTGTEKPYCEICGRKFT-----FSADVQRHIVHTGKPKHLCIDICGRFGNSV 534
 QY 433 SQVAHLRAHV-----LIHTGEKPYCPTCGTRFRLHQLTKS 468
 DB 535 NLKEHEKIHVSDKIYTCDECKSGFMRLKMKHRIHSHTGKPKYNCSTCGKPKAGSGDLQR 594
 QY 469 HVRIHTGEKPYHCDPCGLHFRHKSQRLHLRKHQAATN 507
 DB 595 HVRSHTEKPYCTDCSKNFSRAVLRHKHMKHKAEN 633

RESULT 12

Q6ZSB9 PRELIMINARY; PRT; 765 AA.
 ID Q6ZSB9
 AC Q6ZSB9;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ45653.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Tongue;
 RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
 RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma A., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama M., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Manuho Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AKI27560; BAC87035.1; -.

DR HSSP; P08046; IAIQ.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR011069; Asp_transf_reg_C.
 DR InterPro; IPR000210; BTB_POZ.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00651; BTB; 1.
 DR Pfam; PF00096; zf-C2H2; 7.
 DR ProDom; PD000003; Znf_C2H2; 1.
 DR SMART; SM00325; BTB; 1.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PS00097; BTB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
 SQ SEQUENCE 765 AA; 85076 MW; P0BDF20C2485F590 CRC64;

Query Match 19.3%; Score 539.5; DB 2; Length 765;
 Best Local Similarity 28.5%; Pred. No. 8e-19;
 Matches 169; Conservative 68; Mismatches 224; Indels 131; Gaps 19;

QY 20 HSSDVLGNLRLRGLITDVTLLVGGQPLRAHKAVALIACSGFFYSIFRGRAGVGDVLS 79
 DB 7 HSCILLOLHQRIQGLLDCMLVGVKVRFAKRVNLAAPSOYFRSLFQNSSSQKNDVPH 66
 QY 80 LPGAPEARGFAPLLDFMYTSRLRLSPATAPAVLAATAATYLMQEHVVOACHRFIQAS 137
 DB 67 L-DVKNVSGIGQILDPMYTSRLSHLDLNDNIQVMDLTAQCLOVQNVLSLCHTFLKSATVQP 125
 QY 138 LG-----ISLRPLEAEPPT-----PP-----TAPPGSPRRSEGHDP-PTESRSC 177
 DB 126 PGPMCNSTLSLQSTLTPDATCVISENYPHLLQESADAQONKTLDSEHPHASPVRHH 185
 QY 178 SQGPPSPASDPK--AC-----NWKYKIYVLNSQASQ-----AGSLVGRSSSGQPCQAR 226
 DB 186 SAGEISKQADTSDGCTELPFQPNYYKLRNFYSKYQYHKAAGPQSERVVEQPF--AP 243
 QY 227 LPSGDEASSSSSSSSSSSS-----EGP-----IPQSRLSPTAATVQPK- 267
 DB 244 STSTDLTTVESQCAVSHSECILESPEHLPSNFLAQPVNDSAPHESDATCQPVKQML 303
 QY 268 -----CGAPSTPYL-----LTQAQDT-SGSPSERARPLSGSEFFS 303
 DB 304 KKAHLKLNFLKSQKYAEQVSEPKSDGLTKLESASKNLTLEKASSQSAPEKESEVVS 363
 QY 304 CONCEAVAGCSGLDSLVPQDE-----DKPYKQLCRSSFRYKGNLASHRTVHTGKPY 357
 DB 364 CENFCISETERPEDPAALDQSQTLQSQRYACELCGKFKHPSNLELHKRSHTGKPP 423
 QY 358 HCSICGARPNRPNALKTHRIHSGEKPKECTCGSRFVQVRQPSGFGCKPARGGV--- 414
 DB 424 ECNICGKHFSQAGNLQTLHRRHSGEYPYICEICGRFA-----ASGDVQRH 469
 QY 415 -----GQKGFCSQSRQDLKSPSQVAHLRAHV-----LIH 445
 DB 470 ITHSGEKPHLDCICGRFGNSFNLSNLEKHKHTHTADKIFTCECKSGSPNMOKLVRH 529
 QY 446 TGEKPYCPTCGTRFRLHQLTKSHVRIHTGEKPYHCDPCGLHFRHKSQRLHL 497
 DB 530 TGERPYSCSACGKFCGSGDLRHRHVHTHTGEKPYTCEICNKCFTRSVLRHH 581

RESULT 13

ZP67 HUMAN
 ID ZP67 HUMAN STANDARD; PRT; 539 AA.
 AC Q15156; Q96BP2;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Zinc finger protein 67 homolog (Zfp-67) (Kruppel-related zinc finger protein cKrox) (zinc finger and BTB domain containing protein 15).
 DE Name=ZFP67; Synonyms=ZBTB15;

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC TISSUE=Skin fibroblast;
 RX MEDLINE=98036076; PubMed=9370309; DOI=10.1016/S0378-1119(97)00360-0;
 RA Widom R.L., Culic I., Lee J.Y., Korn J.H.;
 RT "cloning and characterization of hckrox, a transcriptional regulator
 of extracellular matrix gene expression.";
 RL Gene 198:407-420(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
 RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Transcriptional repressor of the collagen COL1A1 and
 COL1A2 genes. May also function as a repressor of fibronectin and
 possibly other extracellular matrix genes.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 BTB/POZ domain.
 CC -1- SIMILARITY: Contains 4 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF007833; AAC51847.1; -
 CC EMBL; BC012070; AAH12070.1; -
 CC HSP; P08047; 1SP2.
 CC Genew; HGNC:18668; ZFP67.
 CC H-InvDB; HIX001022; -
 CC GO; GO:0003704; F:specific RNA polymerase II transcription fa. . . ; TAS.
 CC GO; GO:0007398; P:ectoderm development; TAS.
 CC InterPro; IPR00210; BTB_POZ.
 CC InterPro; IPR007087; Znf_C2H2.
 CC Pfam; PF00651; BTB; 1.
 CC Pfam; PF00036; zf-C2H2; 4.
 CC SMART; SM00225; BTB; 1.
 CC SMART; SM00355; Znf_C2H2; 4.
 CC PROSITE; PS00097; BTB; 1.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 CC DNA-binding; Metal-binding; Nuclear protein; Repeat; Repressor;
 Transcription regulation; Zinc-finger.
 CC DOMAIN 34 115
 CC PRO-rich.
 CC ZN_FING 181 196
 CC C2H2-type 1.
 CC ZN_FING 374 396
 CC C2H2-type 2.

FT ZN_FING 402 424 C2H2-type 3.
 FT ZN_FING 430 454 C2H2-type 4 (atypical).
 FT CONFLICT 437 437 K -> R (in Ref. 1).
 FT CONFLICT 480 480 S -> F (in Ref. 1).
 FT CONFLICT 509 510 GP -> WA (in Ref. 1).
 SQ SEQUENCE 539 AA; 58026 MW; 86B3616504B05B7E CRC64;
 Query Match 18.3%; Score 509.5; DB 1; Length 539;
 Best Local Similarity 28.3%; Pred. No. 1.7e-17;
 Matches 160; Conservative 53; Mismatches 190; Indels 163; Gaps 18;
 QY 1 MGSPPAEGALGVREFTBSSDVLGNELRLGILTVTLVGGOPLAHKAVALIACS 60
 DB 1 MGSP--EDDLIGI--PFPDHSSELLSLNEQRLGHLCDLTIITQGLYETHRAVLACS 56
 QY 61 GFYISIF-----RGRAGVGVVLSLPGGPEARGFAPLLDFMYTSLRL 103
 DB 57 HFYKCLFTGGGAVNGAGSGTGATGAGAGVCELDV-FGPEALG--ALLEFAYTATLT 113
 QY 104 SPATAPAVLAATYLMQEHVQACHRFIOASVEPLGISRLPLEAEPTTP----- 152
 DB 114 SSANMPAVLQARLLLEIPCVIAACMEILQGS-----GL-----EAPSPDEDCERARQY 162
 QY 153 -----PTAPPGSPRSEGHDPDPTESRSCSQGPPSPASDPKACNKKYKIVILNSQASQ 208
 DB 163 LEAFATATAGVPGNSDPPQVPLP-----PPPPPPRPVARRRKRKPAFLQTKGAR 215
 QY 209 AGSLVGRSGGPGCPQARLPISGDEASSSSSSSSSEEGPIPGQSRSLSPTAATVQFKC 268
 DB 216 ANHLVPE--VPTVAHPLTVEEVEVAGRVSSGGS-----GPQDSYSPPTGTASPE 265
 QY 269 GAPASTPYLLTSAQD-----TSGSPSRARPLPGSEFFSCQNC---EAVACSS 315
 DB 266 GPQSYEPYEGEEBELVYPAYGLAQGGP-----PLSPEELGSDDAIDPDLMAYLSS 320
 QY 316 -GLDSLVPG--DEDK-----PYKCOLCRSSFRYKGNLASHRTVHTGKPVHCSICG 363
 DB 321 LHQDNLAPGLDSQDLVLRKRSQMPQECVPCHKIIGHAGKLPRHMRTHTGKPPACEVCG 380
 QY 364 ARFNRPANLKTSHRHSGEKPYKCTCGSRFVQVRQPPSGFGQKPARGGVGQKGFCS 423
 DB 381 VRTRNDKLIHNR-----
 QY 424 QRODLKSPSQVAHLRAHVLHITGKPYCPTCGTFRHLQTLKSHVRIHTGKPYHCDP 483
 DB 395 -----KHTGERPYSCHPCPARFLHSYDLKNHMLHTGDRPYECHL 434
 QY 484 CGLHFRHKSQRLRLHKLQKGAATNTK 509
 DB 435 CHKAFAKEDHLRHLKQNCLEVRTR 460

RESULT 14

Q68DR2 PRELIMINARY; PRT; 573 AA.

AC Q68DR2
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein DKFZp686G01254.
 GN Names=DKFZp686G01254;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RG The German cDNA Consortium;
 RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
 RA Fobo G., Han M., Wiemann S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR749303; CAH18158.1; -


```
FT CONFLICT 268 268 A -> T (in Ref. 2).
FT CONFLICT 328 336 DALTPGLDG -> GRPDRPGW (in Ref. 2).
SQ SEQUENCE 544 AA; 58917 MW; 5A43F10AE2970709 CRC64;

Query Match 17.6%; Score 491.5; DB 1; Length 544;
Best Local Similarity 27.7%; Pred. No. 1.3e-16;
Matches 154; Conservative 56; Mismatches 207; Indels 139; Gaps 14;

QY 1 MGSPAPAEAGALGVYRETRHSSDVILGNELRLRLGILTDVTLVGGQPLRAHKAVLIACS 60
Db 1 MGSP--EDDLIGI--PPDHSSBLLSCNQRQLGHLCDLIRTOGLEIYTHRAVLAACS 56
QY 61 GFFYSIF-----RGRAGYGVVDVLSLPGGPEARGFAPLLDFMYTSRLRL 103
Db 57 HYFKLFTGGGTVMTGGGTASGGAGAGVCELDV-F-GPEALG--ALLEFAYTATLTT 113
QY 104 SPATAPAVLAAATYQWHEVQACHRFIQASYEPLGI-----SLRPLEABPPT 151
Db 114 SSANMPAVLQAAARLLEIPCUIACMETILOGS---GLEAPSPDEDDCERARQYLEAFATA 169
QY 152 PPTAPPSPRRSEGHDPPTESRSCSQPPSPASDPDKACNKKYKIVILNSQASQAGS 211
Db 170 TTTASTSGMNGEDSPQVPLL-----PPPPPPRPVARRRKRKPAFLQTGAKANH 222
QY 212 LVGERSGQPCQARLPFGDEASSSSSSSSSEEGPIPGQSRSLSPSTAATVQFKC--- 268
Db 223 LVPAPTULTHP---LTYEEEMVGRLGNSGGSLGDSYSPPTGAASPAEGPLNYEVFEG 279
QY 269 ----GAPASTPYLTSQAQDTSGSPSERARPLPGE-----PFSCONCEAVAGCS 314
Db 280 EEEEEEMAYPPGYGLAQSNFSLSPSEL-----GSEDEPIDPDLWAYLSSLIHQDALTPGL 334
QY 315 SGLDSLVPGEDDK-PYKQCLCRSSFYKGNLAHRTVHTGKPYHCSICGARFNRPANLK 373
Db 335 DGQDKLVRRRSQMPQECVCHKIIGHAGKLPFRHMRTHGKPFACEVCGVRFTNRNDKLK 394
QY 374 THSRHSGEKPYKCTCGSRFQVRSQPPSGFGKPARGGVGQKGFCSQORQLKSPPS 433
Db 395 IHMR----- 398
QY 434 QVAHLRAHVLHTGKPYPCPTCGTRFRHLQTLKSHVRIHTGKPYHCDPCGLHFRHKSQ 493
Db 399 -----KHTGERPYSCPCPARFLHSYDLAKNMHMLHTGDRPYEHLCHKAFAKEDH 448
QY 494 LRLHLROKHGAATNTK 509
Db 449 LQRHLKGQNCLEVRTR 464
```

Search completed: August 9, 2005, 14:05:16
Job time : 181 secs